

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 08:52:35 ; Search time 10500.1 Seconds  
(without alignments)  
11165.928 Million cell updates/sec

Title: US-10-041-007-1  
Perfect score: 2705  
Sequence: . 1 atatttgcttaacctgcac.....aaaaaaaaaaaaaaaaaaaaa 2705

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_mam.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2705	100.0	2705	8	AF331704	AF331704 Ginkgo bi
2	1091.4	40.3	2861	6	BD227676	BD227676 Syntheses
3	1091.4	40.3	2861	6	AR240711	AR240711 Sequence
4	1091.4	40.3	2861	6	AR266986	AR266986 Sequence
5	1091.4	40.3	2861	6	AR316335	AR316335 Sequence
6	1091.4	40.3	2861	6	AR338478	AR338478 Sequence
7	1091.4	40.3	2861	6	AR429884	AR429884 Sequence
8	1091.4	40.3	2861	6	AGU50768	U50768 Abies grand
9	731.6	27.0	2525	6	AR230308	AR230308 Sequence
10	730.2	27.0	2424	6	BD227671	BD227671 Syntheses
11	730.2	27.0	2424	6	AR222107	AR222107 Sequence
12	730.2	27.0	2424	6	AR230276	AR230276 Sequence
13	730.2	27.0	2424	6	AR240706	AR240706 Sequence
14	730.2	27.0	2424	6	AR266981	AR266981 Sequence
15	730.2	27.0	2424	6	AR316330	AR316330 Sequence
16	730.2	27.0	2424	6	AR338473	AR338473 Sequence
17	730.2	27.0	2424	6	AR429879	AR429879 Sequence
18	730.2	27.0	2424	6	AF006194	AF006194 Abies gra
19	730.2	27.0	2528	6	AR230307	AR230307 Sequence
20	730.2	27.0	2528	6	AR230309	AR230309 Sequence
21	730.2	27.0	2528	6	AF006195	AF006195 Abies gra
22	730.2	27.0	2571	6	AR230287	AR230287 Sequence
23	642.2	23.7	2700	6	AR110171	AR110171 Sequence
24	642.2	23.7	2700	6	BD227670	BD227670 Syntheses
25	642.2	23.7	2700	6	AR240705	AR240705 Sequence
26	642.2	23.7	2700	6	AR266980	AR266980 Sequence
27	642.2	23.7	2700	6	AR316329	AR316329 Sequence
28	642.2	23.7	2700	6	AR338472	AR338472 Sequence
29	642.2	23.7	2700	6	AR382645	AR382645 Sequence
30	642.2	23.7	2700	6	AR429878	AR429878 Sequence
31	642.2	23.7	2700	6	TBU48796	U48796 Taxus brevi
32	627.6	23.2	2706	8	AY007207	AY007207 Taxus chi
33	623	23.0	2643	8	AY424738	AY424738 Taxus bac
34	474.6	17.5	2198	8	AF369918	AF369918 Picea abi
35	472.6	17.5	2150	8	AF369919	AF369919 Picea abi
36	450.4	16.7	2186	6	AR222137	AR222137 Sequence
37	450.4	16.7	2186	6	AF139205	AF139205 Abies gra
38	435	16.1	2100	8	AF543530	AF543530 Pinus tae
39	426.4	15.8	2012	8	AF369920	AF369920 Picea abi
40	418.2	15.5	2186	8	AF461460	AF461460 Picea abi
41	418.2	15.5	2490	8	AB042424	AB042424 Croton gu
42	416.4	15.4	1893	8	AF139206	AF139206 Abies gra
43	416	15.4	1890	6	AR222146	AR222146 Sequence
44	412.2	15.2	2205	6	AR222116	AR222116 Sequence
45	411.2	15.2	1865	6	BD227672	BD227672 Syntheses

ALIGNMENTS

RESULT 1  
AF331704  
LOCUS AF331704 2705 bp mRNA linear PLN 03-OCT-2001  
DEFINITION Ginkgo biloba levodipmaradiene synthase mRNA, complete cds.  
ACCESSION AF331704  
VERSION AF331704.1 GI:15865604  
KEYWORDS  
SOURCE Ginkgo biloba (maidenhair tree)  
ORGANISM Ginkgo biloba  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.  
REFERENCE 1 (bases 1 to 2705)  
AUTHORS Schepmann,H.G., Pang,J. and Matauda,S.P.  
TITLE Cloning and characterization of Ginkgo biloba levodipmaradiene synthase which catalyzes the first committed step in ginkgolide





Db 624 GTACAGAAAGGATTTGAATTTCTTCAGGACACAAAGCTGGAAAGATGAAGCTGAT 683  
 QY 710 AATCAGAGCCAAAGTGGATTCAGAGTCGTGTTCTGTCGATGTAGTGAAGCAAAAGC 769  
 Db 684 AGTCATAGCCAAAGTGGATTTGAATAGTATTTCTGCAATGCTAAAGGAAGCTAAATC 743  
 QY 770 TTGGGATTTGGATTTCTCTTATCAGCTCCCTTTCTATCCCAATCCACAAAGCGCCAG 829  
 Db 744 TTAGGCTTGGATCTGCTTTCAGTTTGGCAATTCCTGAACAATCATCGAAAGCGGAG 803  
 QY 830 AAAAGCTTTCAAAGATTTCCCTCAATGTTTTCATTAACCATCAGACGGCGTTGCTTAC 889  
 Db 804 GCTAAGCTTTAAAGGATTTCCCACTGATGTTCTATGCTTCCCAACAACGTTATTGTAT 863  
 QY 890 TCTCTGGAGGTTTGCAGATGTTGGTACTGGCAAGATCACAATCTTCAATCAAGA 949  
 Db 864 TCTTTGGAGGTTTACAAGAAATAGTAGACTGGCAGAAATATGAATCTTCAATCCAG 923  
 QY 950 GACGATCATTTTAAAGCTCCCTGCTATCTACTGCTTGTGCTTCTCATGCACACTCAAAAC 1009  
 Db 924 GATGATCATTTCTGAGCTCTCGGCACTCAAGGGCTGATTCATGGTACAGGGAAC 983  
 QY 1010 AAACGATGCTCCACTTTCTCAACTTCGTGCTGAGCAAAATTTGGGACATGCTTCTTGC 1069  
 Db 984 AAAAGTGTGATTTCTTGAATTTGTCTTGAAGAAATTCGGAACCATGTGCTTGT 1043  
 QY 1070 CATTACCACTTGATCTATTGAAAGCTTCTGGCTGCTGATACAGCTTTGGGA 1129  
 Db 1044 CACTATCCCTTGATCTATTGAAAGCTTGTGGGCGTTGATACAGTTGAGCGGTAGT 1103  
 QY 1130 ATCGATGCTTATTCAAAGAAAGAAATCAAAGAAATCTCTGGATTTAGTACTG 1189  
 Db 1104 ATCGATGCTTATTCAAAGAGGAGATCAAGGAGCATTTGATTTATGTTTACAGCCATTGG 1163  
 QY 1190 GAGCCGAAAGGCGTGGATGGGCAAGATGCAATCTTCTGATGTCGATGACACT 1249  
 Db 1164 GA---CGAAAGAGGCAATGGATGGGCGAGAGAAATCCTGTTCTGATATTGATGATACA 1220  
 QY 1250 GCATGGGCTTTAGATCTTGAGACTTATGATGATCAATGATCTTCAAGATTTCTGGAG 1309  
 Db 1221 GCATGGGCTTTGAAATCTTGAGATTTACATGATGATCAATGATCTTCAAGATTTTAA 1280  
 QY 1310 AATTTGAGAGACGAGAAAGAGACTTTCTTTGCTTTGCGGCTCAAAAGCAAAATTTGTTG 1369  
 Db 1281 ACATTTAGAGATGAGAAATGGGAGTTCTTTGCTTTCTGCTTCAAAACAGAGAGGAT 1340  
 QY 1370 ACCGATAATCTTAACTTTATAGATGTTCAAGATGTTTTCGCGGAGAAAGATAAG 1429  
 Db 1341 ACAGACATGTTAAACGTCATCTGATGTTTCAATGTTTCAATGTTTTCGCGGAGAAAGCATG 1400  
 QY 1430 GAAGAGCTTAAGACCTTCACTACAAATCATCTCCAAATGCTTCTGCCAAACCAACGCA 1489  
 Db 1401 GAAGAGCAAACTCTGTACCGAAAGGATCTGAGGATGCTCTGGAATTTGGATGCC 1460  
 QY 1490 TTTGATAAGTGGCTGCTCAAGAAAGATTTCTTCTGAGAGGTGAGTATGCTATAAAGTAT 1549  
 Db 1461 TTTGACAAATGGGCTTTTAAAGAAATTTTCGCGGAGAGTATGACACTCAAAATAT 1520  
 QY 1550 CCGTGGCATAGATGATGCAAGATTTGGAGCGAAGAGTTTACATAGAGCAATTTGGATCA 1609  
 Db 1521 CCGTGGCATAGATGATGCAAGATTTGGAGCGTATAGAGCTTATTTGAAACTATTTGGGCCA 1580  
 QY 1610 AATGATGCTCTGGCTGGGAAAGACTGTGTATATAGATGCTATATGTAGCAACGAAATAT 1669  
 Db 1581 GATGATGCTGGCTTGGAAAGAACTGTATATATGATGCTATATGATGCTTCAATGAAAGTAT 1640  
 QY 1670 TTGAGCTGGCCAAATTTGGACTTCAATATGCTGAGAGCTTACCAAAAGGAGACTCAA 1729  
 Db 1641 TTAGAACTAGCGAACTGGACTTCAATATGCTGAGTATGATGCTTATACCAACAGAGCTTCAA 1700  
 QY 1730 CACATTTGCTAGCTGGTGGAGAGAAATCGGATTTCAATGATCTTACATTTTACCCCGCAGCGG 1789

Db 1701 GATCTTGAAGGTTGGTGAATAATCATCCGTTTTCAGGATCTGAATTTTCACTCGTGAGCGT 1760  
 QY 1790 CCTGTGAAATATGATTTTCTCAGTGGCGGTAGTATGTTTGAAGCCAGAAATTCGCTGCTTGT 1849  
 Db 1761 GTGAGGAAATATATTTCTCACCAGCATCTTTATCTTTGAGCCGAGTCTTCTAAGTGC 1820  
 QY 1850 AGAATTTGCTTATGCAAGACTTTCTTCCCTCGAGTTATCTAGACGATCTTTTACAGACC 1909  
 Db 1821 AGAGAGGTTTATACAAAACCTTCCAAATTTCACTGTTATTTTAGATGATCTTTTATGACGCC 1880  
 QY 1910 CACGATCTCTGGATGATCTTAAATTTGTTCTGAAAGCGTCCGAAAGATGGGATATCTCT 1969  
 Db 1881 CATGATCTTTAGACGATCTTAAAGTTTTCAGATCAGTCAAAAGATGGATCTATCA 1940  
 QY 1970 GTGCTGGATAGCGTTTCGGGATTAATCAGTTGAAAGTTTCTCTAGGCGTGTACACACA 2029  
 Db 1941 CTAGTGGG--CCAAATGCCACAACTTGAATGAAATATGTTTGTGGGTTTCTTACAATCT 1997  
 QY 2030 GTGAATGATTTGGAAAGATGGACTCAAGGAAACAGGCGGTGATGCTGCTGGCTATCTT 2089  
 Db 1998 TTTTATGATATAGCAAAAGAGGAGCTGAGAGGCAAGGCGGCGATGTGCTAGGCTACAT 2057  
 QY 2090 CGAAAGTATGGAGGCGTTCCTGCTCATCTATACAAAGAGCGCAATGGTCCGCGACA 2149  
 Db 2058 CAAATGTTTGGAAAGTCCAATTTGAAGCTTACACGAAAGAGCAGAAATGGTCTGAAGCT 2117  
 QY 2150 AAGTATGTCCGACATTCACGCAATATGTGAAATATGCAAAAGTGTCCATGACACTTGGC 2209  
 Db 2118 AAATATGTGCCATCTTCAATGAATACATAGAGATGCGAGTGTGTCAATAGCATTTGGA 2177  
 QY 2210 ACAGTCTGACTTAACTCAATCTTTTCTCAGTGGAGAAATTTCTTCTGATTTTACAG 2269  
 Db 2178 ACAGTCTGTTCTCATTTAGTGTCTTTTCTCAGTGGAGGTTCTTACAGATGAGTACTCTCC 2237  
 QY 2270 CAAATGAGCTTTCGGTCCAAATTTCTGCTATCTTGTCTTTTGACTGGAGCACTAATCAAT 2329  
 Db 2238 AAAATGATCGGAACTTAGATTTCTTCACTCATGGGCTTAAACAGGCGTTTGGTGAAT 2297  
 QY 2330 GACACAAAGCTTACAGCCGAGAGAAACCGTGTGTAATGTTTCCAGCTTACAGTGC 2389  
 Db 2298 GACACAAACCTTATCAGCAGAGAGGTCAAGGTGAGGTGCTTCTGCCATACATGT 2357  
 QY 2390 TACATGAGGAAATTCGAGGTGACAGAGGAAAGCTCTAAGTCAATCTTTTATGTTATC 2449  
 Db 2358 TATATGAGGACCACTTAAATCTCTGAGAGAGGCTCTACACATGCTTATAGTGC 2417  
 QY 2450 ATCGACAAACGCTGAAAGGAAATTTGAATTTGGAGTTGGCCAAACCGAGCAGCAATGCCCA 2509  
 Db 2418 ATGGAATATGCTCGAGAGTTGAATGAGGAGTTTGAA-----TAAACAAATACCG 2471  
 QY 2510 TTGTGTGAGAGACTGCTGTTCAACCTGCAAGTGTATGCTGCTGTTTATATGATC 2569  
 Db 2472 GATATTTACAAAGACTGCTTTTGAATCTGCAAGAAATATGCACTCTTTTATATGCAA 2531  
 QY 2570 AGAGATGCTTTGGTATCTC---TGACAAAGAGATGAAAGACCATGTACGCGAACTCTT 2626  
 Db 2532 GGGATGTTTGCACACTATCATGATATGGAATTTAAAGAGCATGTCAAATAATGCTC 2591  
 QY 2627 TTGATCTCTGCTGGCTGATGATA 2648  
 Db 2592 TTCCAAACCTGCTGATGATA 2613

RESULT 3  
 AR240711  
 LOCUS  
 DEFINITION Sequence 55 from patent US 6468772.  
 ACCESSION AR240711  
 VERSION AR240711.1 GI:27285860  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.  
 linear PAT 20-DEC-2002



## REFERENCE 1 (bases 1 to 2861)

AUTHORS Chappell, J., Manna, K. R., Noel, J. P. and Starks, C. M.

TITLE Methods of making modified polypeptides

JOURNAL Patent: US 6468772-A 55 22-OCT-2002;

FEATURES Location/Qualifiers

source

1..2861

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Query Match 40.3%; Score 1091.4; DB 6; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 3e-291;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

QY 350 GTCGATGAGCAATCAATAAAGATCCAGACATTTGGTAGAGAAATCCAGTGCATGTTT 409  
DB 324 GTTGCAGCATCAGACGAGAGCGTATCGAGACATTAATATCCGAGATTAAAGATATGTTT 383  
QY 410 CAGTCCATGGCGACGCTGAACGAATCCATCTGCATATGATACACGCTGGGGGCAAGA 469  
DB 384 AGATGATGGGCTATGGCGAAGCGAATCCCTCTGCATATGACACTGCTTGGGTAGCAAGG 443  
QY 470 ATTCCGCTCAATTGACGGCTCTGGTGCAACCCCAATTTCCCAACGCTTCAATGGATTCTG 529  
DB 444 ATTCCAGCAGTTGATGGCTCTGACACCCCTCACTTTCTCGAGAGGTTGATGATTTCTT 503  
QY 530 AACATCAATGCGACATGGCTCGTGGGGTAGAGATGCAATTTTTCTGGCGTATGACAGA 589  
DB 504 CAAAATCAGTTGAAGATGGCTCTGGGGTGAAGGATTTCACTTTCTTGGCATATGACAGA 563  
QY 590 GTTTTAAACACTCCGCTGCTCCTCACTCTCAAAATATGGAATGAAGCGGCATTTCAA 649  
DB 564 ATACTGGCTACATTGCGATGATATTAATACCTTACCTCTGGCGTACTGGGGAGACAAA 623  
QY 650 GTGCAGAAAGGGTTGAGTTGTGAGAAACACATGGAAGAAATGAAGACGAAGCTGAC 709  
DB 624 GTACAGAAAGTTATGAAATTTCTTCAAGACACAGCTGGAAGATGGAAGTGAAGCTGAT 683  
QY 710 AATCAAGGCAAGTGGATTCGAGGTGCTGTTTCTGCAATGATGATGAAGCAAAAGC 769  
DB 684 AGTCATAGGCGCAAGTGGATTTGAATAGTATTTCTGCAATGCTAAAGGAAGCTAAATC 743  
QY 770 TTGGGATCGATCTCTTATCACCTCCCTTCACTCTCCCAATCCACCAAGCGCCAG 829  
DB 744 TTAGGCTTGGATCGCTTACGATTTGCCATCTCTGAAACAAATCATCGAAAGCGGGAG 803  
QY 830 AAAAACTTCAAAAGATTCCTTCAATGTTCTTCAATAACCATCAGACGGCGTGTCTCTAC 889  
DB 804 GCTAAGCTTAAAGGATTTCCACTGATGTTCTCTATGCCCCTTCAACAAAGTTATTGTAT 863  
QY 890 TCTCTGAGGGTTTGAAGATGTTGGTGAATGCGAAGATCAAAATCTTCAATCAAGA 949  
DB 864 TCTTTGGAAGTTTCAAGAAATAGTAGACTGGCAGAAAATATGAACCTTCAATCCAAG 923  
QY 950 GACGATCATTTTAAGCTCCCTGCACTCTACTGCTTGTGCTTTCATGCACTCAAAAC 1009  
DB 924 GATGATCATTTTCAAGCTCTCGGATCTACAGCGGCTGTATTCTATGCGTACAGGGAAC 983  
QY 1010 AAACGATGCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGCGCACTACGTTCTTGC 1069  
DB 984 AAAGAGTGTGATTTCTTGAATTTTGTCTGAGAAATTCGAAACCAATGTGCTTGT 1043  
QY 1070 CATTACCCATTTGATCTATTGTAACGCTCTGGGCTGTGATACAGTGTGAACGCTTGGGA 1129  
DB 1044 CACTATCCGCTGATCTATTGAACGTTTGTGGGCGTTGATACAGTGTGAGCGGCTAGGT 1103  
QY 1130 ATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGAAATACGTTTATAGTACTGG 1189  
DB 1104 ATCGATCGCTATTTCAGAAAGGAGATCAAGGAAAGCATTTGATATGTTTACAGCCATTGG 1163  
QY 1190 GACGCGAAAGAGCGGTGGGATGGGCAAGATGCAATCTTATCTCTGATGTGATGACACT 1249  
DB 1164 GA---CGAAAGAGGCATTTGATGGGCGAGAGAGAAATCTCTTCTGATATTGATGATACA 1220

QY 1250 GCCATGGGCTTTAGATCTCTGAGACTTCTATGGATACAATGTATCTTTCAGATGTTCTGGAG 1309  
DB 1221 GCCATGGGCTTTGGAATCTTGGATATACATGGATACAATGTATCTTTCAGATGTTTAAAAA 1280  
QY 1310 AATTTTCAGAGACGAGAAAGGAGACTTCTTTTGTCTTTCGCGGTCAAAACGCAAAATTTGGTGTG 1369  
DB 1281 ACATTTAGAGATGAGAAATGGGAGTTCTTTTGTCTTCTGGGTCAAAACAGAGAGGAGTT 1340  
QY 1370 ACCGATAATCTTTAACTTTATAGATTTCAAGTATGTTTTCGGGAGAAAAGATATG 1429  
DB 1341 ACAGACATGTTAAACGTCATCTGTTTCAATGTTTCAATGTTTTCATTTCCGGGAGAAACGATCATG 1400  
QY 1430 GAAGAGCTTAAGACCTTCACTACAAATCATCTCCAAATATGCTCTTTCGCAAAAACAGCA 1489  
DB 1401 GAAGAGCAAAAACCTCTGTATCCGAAAGGTATCTGAGGAATGCTCTGGAATATGTTGGATGCC 1460  
QY 1490 TTTGATTAAGTGGGCTCTCAAGAGAGGATCTTCTGAGAGGTGAGTATGCTATATAAGTAT 1549  
DB 1461 TTTGCAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGAGTATGCACTCAATAT 1520  
QY 1550 CCGTGCAATAGAAATATGCAAGATTTGGAGGCAAGAGTTTACATAGAGCAATTTGGATCA 1609  
DB 1521 CCTGGCATAAGAGATATGCCAAGTTTGGAGGCTTAGAAGCTATATTGAAAACATATGGGCCA 1580  
QY 1610 AATGATGCTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAAAAATAT 1669  
DB 1581 GATGATGTGTGGCTTGGAAAAAATCTGTATATATGATGCCATACATTTTTCGAATGAAAAGTAT 1640  
QY 1670 TTGGAGCTGGCCAAATTTGGACTTCAATATGTTGTCAGGCTTACACCAAAAGAGAGACTCAA 1729  
DB 1641 TTAGAACTAGCGAACTGGAATCTTCAATAGGTGAGTCTATACACCAAAAGAGAGCTTCAA 1700  
QY 1730 CACATTTGTCAGCTGGTGGAGAGAAATCGGATTTCAATGATCTTACATTTCACTCCGCGACGG 1789  
DB 1701 GATCTTTGGAAGGTGGTGGAAATCATCCGCTTTTCAAGGATCTGAATTTCACTCGTGAGCGT 1760  
QY 1790 CTTGTGGAATATGATTTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCTGCTGT 1849  
DB 1761 GTGACGGAATATATTTCTCACCGGCATCTCTTATCTTTGAGCCGAGTTTCTTAAGTGC 1820  
QY 1850 AGAATTTGCTTATGCCAAGACTTCTTGCCTCGCAAGTTTATTTCTAGACGATCTTTACGACACC 1909  
DB 1821 AGAGAGTTTATACAAAACCTTCCAAATTTCACTGTTATTTTATAGTATCTTTATAGCGCC 1880  
QY 1910 CACGGAATCTCTGATGATCTTAAATTTGTTCTCTGAAGCGGTCCGAAAGATGGGATATCTCT 1969  
DB 1881 CATGGATCTTAGACGATCTTAAGTTGTTTACAGAAATCAGTCAAAAGATGGGATCTATCA 1940  
QY 1970 GTGCTGGATAGCGTTCCGGATTAATCAGTTGAAAGTTTGTCTTCTAGGGCTGTACACACA 2029  
DB 1941 CTAGTGGG---CCAAATGGCCACAAACAAATGAAAATATGTTTGTGGGTTTCTACAACT 1997  
QY 2030 GTCAATGGATTTGGAAAAGATGAGCTCAAGGAACAGGCGGTGATGCTGGGCTATCTT 2089  
DB 1998 TTTAATGATATAGCAAAAGAGACGTGAGAGCGGCGCGATGTGTGAGGCTACAT 2057  
QY 2090 CGAAAAGTATGGAGGGCTTGTCTGATCGTATATACCAAAAGAACCGAAATGGTGGGAGCA 2149  
DB 2058 CAAAATGTTTGGAAAGTCCAACTTGAAGCTTACCGAAAGAGACAGAAATGCTCTGAAGCT 2117  
QY 2150 AAGTATGTCGCGACATTTCAACGAATATGGAAGAAATGCGAAAGATGTCATAGCACTTGG 2209  
DB 2118 AAATATGTGCCATCTCTCAATGAATACATAGAGAAATGCGAGTGTGTCAATPAGCATTTGGGA 2177  
QY 2210 ACAGTGTACTAACTCAATCTTTTCTGAGTAAATTTCTTCTGAGTAACTTCTTACAG 2269  
DB 2178 ACAGTGTCTCATTAAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTACTCTCC 2237  
QY 2270 CAAGTAGACCTTGGGTCCTCAAAATTTCTGCTCTTGTCTTTGACTGGAGCACTAACTCAAT 2329  
DB 2238 AAAATTTGATCGCGAATCTAGATTTCTTCACTCATGGGCTTAAACGGGCGTTTGGTCAAT 2297

QY 2330 GACACCAAGACTTACACAGCCGAGAGAAACCGTGGTGAATGGTTTCCAGCGTACAGTGC 2389  
Db 2298 GACACCAAACTTATCAGGCAGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATACAATGT 2357  
QY 2390 TACATGAGGGAATATCCCGAGGTGCACAGAGGAAGACTCTAAGTCATGTTTATGGTATC 2449  
Db 2358 TATATGAAGGACCATCTAAATATCTCTGAAGAAGAGCTCTACAACATGTCTATAGTGTG 2417  
QY 2450 ATCGACACGCACTGAGGAATGGAATGGAGTTGGCCCAACCCAGCGCAATGCCCA 2509  
Db 2418 ATGGAAATGCCCCGAGAGTTGAATAGGAGTTTGTGA-----TACAAAATACGG 2471  
QY 2510 TTGTGTGTGAGAGACTGTCTTCAACACTGCAAGAGTGAATGAGTGTCTTTTATATGTAC 2569  
Db 2472 GATATTACAAAGACTGTGTTTGAACCTGCAAGATAATGCAACTCTTTTATATGCAA 2531  
QY 2570 AGAGATGGCTTTGGTATCTC-----TGACAAAGAGATGAAGACCATGTGACCGAACTCTT 2626  
Db 2532 GGGGATGGTTTGACACTATCACATGATATGGAATTAAGAGCATGTCAAAAATATGCCCTC 2591  
QY 2627 TTCGATCTGTGCGGTAGCATA 2648  
Db 2592 TTCCAACAGTTGCCCTAGATTA 2613

RESULT 4  
AR266986  
LOCUS AR266986 2861 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 55 from patent US 6495354.  
ACCESSION AR266986  
VERSION AR266986.1 GI:29696441  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 2861)  
AUTHORS Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.  
TITLE Syntheses  
JOURNAL Patent: US 6495354-A 55 17-DEC-2002;  
FEATURES Location/Qualifiers  
source  
1..2861  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 40.3%; Score 1091.4; DB 6; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 3e-291;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

QY 350 GTCGATGAGACAATCAATAAAGAATCCAGACATTCGTGAAGAAATCCAGTGCATGTTT 409  
Db 324 GTTGCAGCATCAGACGAGAAGCGTATCGAGACATTAATCCGAGATTAAAGAATATGTTT 383  
QY 410 CAGTCATCGGCGACGCTGAACGATCCATCTGCATATGATACAGCTTGGGTGGCAAGA 469  
Db 384 AGATGTATGGGCTATGCGGAACGAATCCCTCTGCATATGACACTCTGGGTAGCAAGG 443  
QY 470 ATTCCGTCAATTGACGGCTCTGGTGACACCCCAATTTCCCAACGCTTCAATGGATTCTG 529  
Db 444 ATTCCAGCAGTTGATGGCTCTGACACCCCTCACCTTCTCGAGACGGTTGNAATGATCTT 503  
QY 530 AACAAATCAATGCGCATGCTCGTGGGTGAGGATGCAATTTTCTGGCGTATGACAGA 589  
Db 504 CAAAATCAGTTGAAGATGGCTCTGGGTGAAAGATTCTACTTCTTGGCATATGACAGA 563  
QY 590 GTTTTAAACACTCTCGCTGCCTCTCTCACTCTCAAAATATGGAATGAAGCGCATTTCAA 649  
Db 564 ATACTGGCTACACTTGCATGTATATATACCTTACCTCTGGCGTACTTGGGGAGACAAA 623  
QY 650 GTGACAGAAAGGGTGTGAGTTTGTGAGAAAACACATGGAAGAAATGAAGGACGAAGCTGAC 709  
Db 624 GTACAGAAAGGTATTGAAATCTTCAGGACACAGCTGGAAAGATGGAAGATGAAGCTGAT 683

QY 710 AATCACAGCCCAAGTGGATTCGAGGTCTGTGTTCTGCAATGTTAGATGAACCAAAAGC 769  
Db 684 AGTCATAGCCCAAGTGGATTTGAAATAGTATTTCTGCAATGCTAAAGGAAGCTAAATC 743  
QY 770 TTGGATTTGGATCTCTCCCTATACACTCCCTTTCATCTCCCAATCCACCAAGCGCCAG 829  
Db 744 TTAGCTTTGGATCTGCTCTTACGATTTGCAATTTCTTGAACCAATCATCGAAAGCGGGAG 803  
QY 830 AAAAAGCTTTCAAAAGATTTCCCTCAATGTCTTCTATAAACCATCAGACGCGTGTCTAC 889  
Db 804 GCTAAGCTTTAAAGATTTCCCACTGATGTCTCTATGCCCTTCCAAACAGCTTATTGTAT 863  
QY 890 TCTCTGAGGGTTTCAAGATGTGTGGACTGCGCAAGAGATCACAAATCTTCAATCAAGA 949  
Db 864 TCTTTGGAAGGTTTACAAGAAATAGTAGACTGGCAGAAAATAATGAATCTTCAATCCAA 923  
QY 950 GACGATCATTTTAAAGCTCCCTGCACTACTGCTGTGCTCTCATGCACACTCAAAAC 1009  
Db 924 GATGATCATTTTCTCAGCTCTCCGCACTTACAGCGGTGTATTTATGCGGTACAGGGAAC 983  
QY 1010 AAAAGATGCTCCACTTTCTCAACTTCGTGCTCAGCAAAATTTGGCGACTACGTTCTTGC 1069  
Db 984 AAAAGATGCTTGGATTTCTTGAACCTTTGTCTTGAAGAAATTCGAAACCATGTGCTTGT 1043  
QY 1070 CATTTACCACTGATCTATTTGAACGCTCTGGGCTGTGCATACAGTTGAAACGCTTGGGA 1129  
Db 1044 CACTATCCGCTTGATCTATTTGAACGCTTTGTGGGCGGTGATACAGTTGAGCGGTAGGT 1103  
QY 1130 ATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATAGCTTTATAGGTACTGG 1189  
Db 1104 ATCGATCGCTATTTCAAGAGGAGATCAAGAGCATTTGGATTTATGTTTACAGCCATTGG 1163  
QY 1190 GACGCGAAAGAGCGGTGGGATGGCAAGATCCAAATCTTATTCCTGATCTCGATGACACT 1249  
Db 1164 GA---CGAAGAGGCATTGGATGGCGAGAGAAATCTGTGTTCTCTGATATTGATGATACA 1220  
QY 1250 GCCATGGGCTTTAGAAATCTCGAGACTTCATGATACAAATGATCTTCAGATGTTCTGGAG 1309  
Db 1221 GCCATGGGCTTTCGAATCTTGAGATTACATGATACAAATGATCTTCAGATGTTTAAAA 1280  
QY 1310 AATTTTCAGAGACGAGAAAGGAGACTTTCTTCTGCTTGGCGGTCAACGCAAAATTTGTGTG 1369  
Db 1281 ACATTTAGAGATGAGAAATGGGAGTTCTTTTCTTCTTGGGTCAAAACACAGAGAGAGTT 1340  
QY 1370 ACCGATAATCTTAACTTTATAGATGTTTCAAGATATGTTTTCGGGAGAAAGATATG 1429  
Db 1341 ACAGACATGTTAAACGTCATCAATCGTTGTTTCAATGTTTCAATTTCCGGGAGAAACGATCATG 1400  
QY 1430 GAAGAGCTTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTCCCAAAAACAAACGCA 1489  
Db 1401 GAAGAAGCAAAACTCTGTACCCGAAAGGTATCTGAGGAATGCTCTGAAATATGGGATGCC 1460  
QY 1490 TTTTGATAAGTGGGCTGTCAAGAAAGGATCTTCTGGAGAGGTGGAGTATGCTATAAAGTAT 1549  
Db 1461 TTTGACAAATGGCTTTTAAAGAAATAATTCGGGAGAGGTAGATATGCACTCAAAATAT 1520  
QY 1550 CCGTGCAATAGAAGTATGCCAAGATTTGAGGCGAAGAGTTACATAGAGCAATTTGGATCA 1609  
Db 1521 CCTGCGCATAGAAGATGCAAGGTTTGGAGGCTAGAAGCTATATTGAAAACATATGGGCCA 1580  
QY 1610 AATGATGCTGCTGGGGAAGACTGTGATAAGATGCTATATGTGAGCAACGAAAATAT 1669  
Db 1581 GATGATGTGGCTTGGAAAAAACTGTATATATGATGCCATACATTTTCAATGAAAGTAT 1640  
QY 1670 TTGGAGCTGGCCAAATTTGACTTCAATATGGTGCAGGCTTACACAAAGAGGAGACTCAA 1729  
Db 1641 TTAGAACTAGCGAAACCTGCACTTCAATAAGGTGCACTATATACACAAACAGAGCTTCAA 1700  
QY 1730 CACATTTGCTGCTGGTGGAGAGAAATCGGATTTCAATGATCTTACATTTACCCGCCGCGG 1789  
Db 1701 GATCTTTCGAAGGTTGGGAAATCATCCGGTTTCAGGATCTGAATTTCACTCTGTGAGCGT 1760  
QY 1790 CCTGTGGAATATGATTCTTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCGCTGCTTGT 1849

Db 1761 GTGACGGAAATATATTTCTCACCGGATCTCTTTATCTTTGAGCCGAGTTTCTTAAGTGC 1820  
Qy 1850 AGAATGCTATGCCAAGACTTCTTGCTCGCAGTATCTTAGCAGTCTTTTACGACACC 1909  
Db 1821 AGAGAGGTTTATACAAAACCTTCCAAATTTCACTGTTATTTTAGATGATCTTTATGACGCC 1880  
Qy 1910 CACGATCTCTGGATGATCTTAAATTTGTTCTGAAAGGGTCCGAAAGTGGATATCTCT 1969  
Db 1881 CATGGATCTTTAGACGATCTTAAAGTTGTTCAAGAAATCAGTCAAAAGATGGATCTATCA 1940  
Qy 1970 GTGCTGGATACGTTTGGGATAATCAGTTGAAAGTTTGTCTTCTAGGCTGTCAACACA 2029  
Db 1941 CTAGTGGG---CBAATGCCACAATAATGAAATATGTTTGTGGGTTTCTCAATACT 1997  
Qy 2030 GTGAATGGAATTTGGAAGAGTGGACTCAAGGAACAAGCCCGTGTGCTGGGCTATCTT 2089  
Db 1998 TTTAATGATATAGCAAAAGAGGCGTGAGAGGCAAGGGCGCGATGTGCTAGGCTACAT 2057  
Qy 2090 CGAAAGTATGGAGGGTCTGCTCGCATCTATACCAAGAGCGGAATGGTCCGACGA 2149  
Db 2058 CAAAATGTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGAGCAGAAATGGTCTGAAGCT 2117  
Qy 2150 AAGTATGTGCGACATTCACGAATATGTGAAATATGCCAAAGTGTCCATAGCACTTGGC 2209  
Db 2118 AAATATGTGCCATCTTCAATAGTCTCTTTTCACTGGGAGGTTCTTACAGATGAAGTACTCC 2177  
Qy 2210 ACAGTGTCTAATAACTCAATCTTTTCACTGGAGAAATTTACTTCTGATTAATTTTACAG 2269  
Db 2178 ACAGTGTCTCTATTAGTCTCTTTTCACTGGGAGGTTCTTACAGATGAAGTACTCTCC 2237  
Qy 2270 CAAGTAGACCTTCCGGTCCAAATTTCTGATCTTGTGCTTTGACTGGAAGCACTAATCAAT 2329  
Db 2238 AAAATTTGATCGCGAATCTAGATTTCTTCAACTCATGCGCTTAAACAGGCGGTTTGGTGAAT 2297  
Qy 2330 GACACCAAGCTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGGTTACAGTGC 2389  
Db 2298 GACACCAAACTTATCAGGCGAGAGAGGTTCAAGGTGAGGTGGGTTCTTGCACATCAATGT 2357  
Qy 2390 TACATGAGGGAATATCCGAGTGACAGAGAGAAAGCTCTTAAGTCTATGTTTATGGTATC 2449  
Db 2358 TATATGAAGGACCATCTTAATCTCTGAAGAGAGCTCTACAACTGCTATAGTATC 2417  
Qy 2450 ATCGAACAGCACTGAAGGAATGAAATGGAGTTGGCCAAACCCAGGAGCAATGCCCCA 2509  
Db 2418 ATGGAATAATGCCCTCGAAGAGTTGAATAGGAGTTTGTGA-----TAACAAAATACCG 2471  
Qy 2510 TTGTGTGTGAAGACTGCTCTTCAACACTGCAAGAGTGTGAGCTGTTTATATATCTAC 2569  
Db 2472 GATATTTACAAAGACTGGTTTGTGAACCTGCAAGAAATATGCAACTCTTTTATATGCA 2531  
Qy 2570 AGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAAGACCATGTGACGCCAACTCTT 2626  
Db 2532 GGGATGGTTTGACACTATCATATGATATGAAATTAAGAGCATGTCAAAAATGGCTC 2591  
Qy 2627 TTGATCTCTGGCGGTAGCAT 2648  
Db 2592 TTCCAACCAAGTTGCTAGATTA 2613

RESULT 5  
AR316335  
LOCUS AR316335 2861 bp DNA linear PAT 12-JUN-2003  
DEFINITION Sequence 55 from patent US 6559297.  
ACCESSION AR316335  
VERSION AR316335.1 GI:31711070  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2861)  
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.  
TITLE Syntheses

JOURNAL Patent: US 6559297-A 55 06-MAY-2003;  
FEATURES Location/Qualifiers  
source 1. 2861  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 40.3%; Score 1091.4; DB 6; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 3e-291;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;  
Qy 350 GTGATGAGCAATCAATAAAGAACCCACACATCTGGTGAAGGAAATCCAGTGCATCTTT 409  
Db 324 GTTGCAGCATCAGACGAGAAAGCGTATCGACACATTAATATCCGAGATTAAGAATATGTT 383  
Qy 410 CAGTCCATGCGGCGCGTGAACGAATCCATCTGCATATGATACAGCTTGGGTGGCAAGA 469  
Db 384 AGATGATGCGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTTGGTAGCAAG 443  
Qy 470 ATTCCGCTCAATGACGGCTCTGGTGCAACCCCAATTTCCCAACGCTTCAATGATTTCTG 529  
Db 444 ATTCCAGCAGTTGATGGCTCTGACAAACCTTCACTTTCTCGAGACGGTTGAATGATTTCT 503  
Qy 530 AACAACTCACTGCGAGTGGCTCTGGTGCAACCCCAATTTCCCAACGCTTCAATGATTTCTG 589  
Db 504 CAAAATCAGTTGAAGATGGGTCTTGGGGTGAAGGATTTCTACTTCTTGGCATATGACAGA 563  
Qy 590 GTTTTAAACACTCTCGCTCCCTCCTCACTCTCAAAATATGGAATGAAGGCGACATTTCAA 649  
Db 564 ATACTGGCTACACTTGCATGATATTTATACCTTACCTCTGGGCTACTGGGGAGACAA 623  
Qy 650 GTGCAAGAGGGTTGAGTTTGTGAGAAACACATGGAAGAAATGAAGAGCAAGCTGAC 709  
Db 624 GTACAGAAAGGTATGAAATTTCTTTCAGGACACAAAGCTGGAAGATGGAAGTGAAGCTGAT 683  
Qy 710 AATCAGAGGCAAGTGGATTCGAGTCTGTTTCTGCAATGTTAGATGAAGCAAAAGC 769  
Db 684 AGTCATAGGCAAGTGGATTTGAAATAGTATTTCTGCAATGCTTAAAGGAGCTTAAATC 743  
Qy 770 TTGGGATTTGATCTTCTTATCACCTCCCTTCACTCTCCCAATTCACCAAAAGCGCGAG 829  
Db 744 TTAGGCTTGGATCTGCTTACGATTTGCCATCTCTGAAACAAATCATCGAAGGCGGAG 803  
Qy 830 AAAAAGCTTCAAAAGATTTCCCTCAATGTTTCTTCAATCAATCAGACGGGCTGCTCTAC 889  
Db 804 GCTAAGCTTAAAGAGATTTCCCACTGATGTTCTCTATGCCCCCTTCCAAACAAGTTATTTAT 863  
Qy 890 TCTCTGGAGGTTTGAAGATGTTGGTGGAGTGGTGGAGAGATCAAAATCTTCAATCAAGA 949  
Db 864 TCTTTGGAAGGTTTACAAAGAAATAGTAGACTGGCAGAAATATGAACACTTCAATCCAAG 923  
Qy 950 GAGCGATCATTTTAAAGCTCCCTGATCTACTGCTTGTCTTCTCATGACACACTCAAAAC 1009  
Db 924 GATGGATCATTTCTCAGCTCTCGGATCTTACAGCGGCTGTATTCATGCGTACAGGAAC 983  
Qy 1010 AAACGATGCTCCACTTTTCAACTTGTGCTCAGCAAAATTTGGCGACTAGTTCCTTGC 1069  
Db 984 AAAAGTGTCTGGATTTCTTGAACCTTGTCTTGAAGAAATTCGAAACCATGTGCTTGT 1043  
Qy 1070 CATTAACCATTTGATCTTATTTGAACGCTCTGGGCTGTGATACAGTTGAACGCTTGGGA 1129  
Db 1044 CACTATCCGCTTGTATCTTATTTGAACGTTTGTGGGCGGTTGATACAGTTGAGCGGCTAGGT 1103  
Qy 1130 ATCGATCGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGATTTACGTTTATAGTACTGG 1189  
Db 1104 ATCGATCGCTTATTTCAAGAGGAGATTCAGAAAGCATTTGATTTATGTTACGCCATGG 1163  
Qy 1190 GACGCCGAAAGAGCGGTGGATGGGCAAGATGCAATCTTATCTGATGTGATGACACT 1249  
Db 1164 GA---CGAAAGAGGCATTGGATGGGCGAGAGAGATCTCTGTTCTTGATAATTGATGATCA 1220  
Qy 1250 GCCATGGGCTTTAGATCTCTGAGACTTTCATGATACAAATGATATCTTCAGATGTTCTGGAG 1309

Db 1221 GCCATGGCCCTCGAATCTTGAGATTACATGATACATATGATCTCTCAGATGTTTAAAA 1280  
Qy 1310 AATTTACAGACGAGAAAGGAGACTTCTTTTGTCTTTGGCCGCTCAAAACCAATTTGGTGTG 1369  
Db 1281 ACATTTAGAGATGAGAAAGGGAGTCTTTTGTCTTTGGGTCAACACAGAGAGGAGTT 1340  
Qy 1370 ACCGATATCTTAACCTTTATAGATGTTTACAAAGTATGTTTCCGGGAGAAAGATAATG 1429  
Db 1341 ACAGACATGTTAAACGTCATCGTTTGTTCACATGTTTCATTTCCGGGAGAAACGATCATG 1400  
Qy 1430 GAAAGAGCTAAGACCTTCACTACAAATCATCTCCAAAATGCTTTGCCAAAACCAACGCA 1489  
Db 1401 GAAAGACCAAACTCTGTACCGAAAGGTATCTGAGGAATGCTCTGAAAATGTGATGCC 1460  
Qy 1490 TTTGATAGTGGGTCTGCMAGAGATCTTCTGAGAGGTGAGTATGCTATAAGTAT 1549  
Db 1461 TTTGACAAATGGGCTTTTAAAAAGAAATATTCGGGAGAGGTAGAGTATGCACTCAATAT 1520  
Qy 1550 CCGTGGCATAGAGTATGCCAAGATTGGAGGCAAGAGTTTACATAGAGCAATTTTGGATCA 1609  
Db 1521 CCTGGCATAGAGTATGCCAAGTTTGGAGGCTAGAGCTATATTGAAAATCTATGGGCA 1580  
Qy 1610 AATGATGCTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTGAGCAACGAAAATAT 1669  
Db 1581 GATGATGCTGGCTTGGAAAACCTGTATATATGATGATGCCATACATTTTCAATGAAAAGTAT 1640  
Qy 1670 TTGGAGCTGGCCAAATTTGACACTTCAATATGTTGAGGCTTACACCAAGAGAGACTCAA 1729  
Db 1641 TTAGAACTAGCGAAATGCACTTCAATAAGGTGCACTTATACCAACCAAGAGCTTCA 1700  
Qy 1730 CACATTTGCTGGTGGGAGAGAAATCGGGATTCATATGATCTTACATTCACCCCGCAGGG 1789  
Db 1701 GATCTTGAAGGTGGTGGAAATCATCCGGTTTCAAGGATCTGAATTTTCACTGCTGAGCGT 1760  
Qy 1790 CTTGTGGAATATGATTTCTCAGTGGCGGTATGATGTTTGGAGCAAAATCTGCTGCTT 1849  
Db 1761 GTGACGGAATATATTTCTCACGGGATCTCTTATCTTTTGGCCGAGTTTCTTAAGTGC 1820  
Qy 1850 AGAATTTGCTATGCCAAGACTTCTTCCCTCGCAGTTTATTTCTAGACGATCTTTACGACAC 1909  
Db 1821 AGAGAGTTTATACAAAATCTTCAATTTTCACTGTTTATTTAGATGATCTTTATGAGCC 1880  
Qy 1910 CACGATCTCTGGATGATCTTAATTTGTTCTTGAAGCGGTCCGAAGATGGGATATCTCT 1969  
Db 1881 CATGGATCTTTAGACGATCTTAAGTTGTTTACAGAAATCAGTCAAAAGATGGGATCTATCA 1940  
Qy 1970 GTCTGTAGAGCTTGGGATATCAGTTGAAGTTTGTCTTCTAGGGCTGTACACACACA 2029  
Db 1941 CTAGTGA---CCAAATGCCACAAATGAAATATGTTTGTGGGTTTCTACATATCT 1997  
Qy 2030 GTGAATGGATTTGGAAAAGATGGACTCAAGGAACAAGCGCGTATGCTGGGCTATCTT 2089  
Db 1998 TTTAATGATATAGCAAAAGAGCGTGTGAGAGCGGGCGCATGTGCTAGGCTACAT 2057  
Qy 2090 CGAAAAGTATGGAGGCTTGTGCGATCGTATACCAAGAACGGAATGGTGGGAGCA 2149  
Db 2058 CAAAAATTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGACGAATGGTCTGAAGCT 2117  
Qy 2150 AAGTATGGCGGACATTCACGATATGTGMAAATGCCAAAGTGTCCATAGCACTTGGC 2209  
Db 2118 AAATATGTGCCATCTTCAATGAATACATAGAAATCGGAGTGTGTCAATAGCATTTGGGA 2177  
Qy 2210 ACAGTCTACTAACTCAATCTTTTTCACCTGGAGAAATTAATCTCTGATTAATTTTACAG 2269  
Db 2178 ACAGTCTTCTCAATAGTCTTTTCACTGGGAGGTTTCTTACAGATGAAGTACTCTCC 2237  
Qy 2270 CAAGTAGACCTTCGGTCCAAATTTCTGCATCTTGTGCTTTTGTGCTGACGACCACTAAAT 2329  
Db 2238 AAAATTGATCGGAATCTAGATTTCTTCAACTCATGGGCTTAAACGGGCGTTTGGTGAAT 2297  
Qy 2330 GACACCAAGACTTACACAGGCGGAGAGAAACCGTGTGAATTTGGTTTCCAGCGTACAGTGC 2389  
Db 2298 GACACCAAACTTATCAGGACAGAGAGGTTCAAGGTGAGGGTCTTCTGCCATACAATGT 2357

Qy 2390 TACATGAGGAAAATCCGAGTGCACAGAGAAAGAACTCTTAAGTATGTTTATGTTATC 2449  
Db 2358 TATATGAGGACCACTCTAAAATCTCTGAAGAGAGAGCTCTTACAACTATGCTATAGTGC 2417  
Qy 2450 ATCGAACCGCACTGAAGGATTTGAATTTGGAGTTGGCCCAACCCAGCGAGCAATGCCCA 2509  
Db 2418 ATGAAAATGCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAAATACCG 2471  
Qy 2510 TTGTGTGTGAGAGACTGCTTCAACACTGCAAGAGTGTGACAGCTGTTTATATGTPAC 2569  
Db 2472 GATATTTACAAAGACTGGTTTTTGAACCTGCAAGATAATGCAACTCTTTTATATGCAA 2531  
Qy 2570 AGAGATGCTTTTGGTATCTC---TGACAAAGAGATGAAAGACCATGTGACGCGAATCTT 2626  
Db 2532 GGGGATGGTTTGACACTATCATCATGATATGGAATTAAGAGCATGTCAAAAATTTGCCTC 2591  
Qy 2627 TTGATCTGTGGCGTAGCATA 2648  
Db 2592 TTCAACCCAGTTGCTAGATTA 2613

RESULT 6  
AR338478 LOCUS 2861 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 55 from patent US 6569656.  
ACCESSION AR338478  
VERSION AR338478.1 GI:33725255  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2861)  
AUTHORS Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.  
TITLE Syntheses  
JOURNAL Patent: US 6569656-A 55 27-MAY-2003;  
FEATURES  
source  
1..2861  
Location/Qualifiers  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 40.3%; Score 1091.4; DB 6; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 3e-291;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

Qy 350 GTCCATGAGACAATCAATAAAGATCCAGACATTTGGTGAAGAAATCCAGTGCATGTTT 409  
Db 324 GTTCGAGCATCAGACGAGAAAGCGGTATCGAGACATTAATATCCGAGATTAAGAAATATGTTT 383  
Qy 410 CAGTCCATGGCGCGCGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTGGCAAGA 469  
Db 384 AGATGTATGGGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTTGGGTAGCAAG 443  
Qy 470 ATTCCGCTCAATTGACGGCTCTGGTGCAACCCCAATTTCCCAACCGTTCATATGGATCTG 529  
Db 444 ATTCCAGCAGTTGATGGCTCTGACAACTCTCACTTTCTCGAGACGGTTGAATGGATCTT 503  
Qy 530 AACAAATCAATGCGACATGCTCGTGGGTGAGGAGTGCATTTTCTGGGTATGACAGA 589  
Db 504 CAAAATCAGTTGAAGATGGGCTTTGGGGTGAGGATTTCTACTTTCTTGGCATATGACAGA 563  
Qy 590 GTTTTAAACACTCTCGCCCTCGCTCTCACTCTCAAAATATGGAATTAAGGCGCACTTCAA 649  
Db 564 ATACTGGCTACACTTGCATGTATTAATCCCTTACCTCTGGCGTACTGGGGAGACAA 623  
Qy 650 GTGCAGAAAGGGGTGAGTTTGTGAGAAACACATCGAAGAAATGAAGACGAGCTGAC 709  
Db 624 GTACAGAAAGGTATGAATTTCTTTCAGGACACAAGCTGGAAAGATGGAAGATGAAGCTGAT 683  
Qy 710 AATCAGAGCGCAAGTGGATTCGAGGTGCTGTTTCTCGAATGTTAGATGAAGCAAAAGC 769  
Db 684 AGTCATAGGCGCAAGTGGATTTGAATAGTATTTTCTGCAATGCTTAAGGAAGCTTAAATC 743



ORIGIN				/organism="unknown" /mol_type="genomic DNA"			
Query Match				40.3%; Score 1091.4; DB 6; Length 2861;			
Best Local Similarity				68.4%; Pred. No. 3e-291;			
Matches 1574; Conservative				0; Mismatches 713; Indels 15; Gaps 4;			
QY	350	GTGATGAGACAATCAATAAAGATCCAGACATTTGGTGAGGAAATCCAGTGCATGTTT	409	Db	1281	ACATTTAGAGATGAGAATGGGGAGTCTCTTTGCTTTCTGGGTCAACACACAGAGGAGTT	1340
Db	324	GTTCGAGCATCAGACGAGAGCGTATCGAGACATTAATATCCGAGATTAAGAATATGTTT	383	QY	1370	ACCGATATCTTAACCTTTATAGATGTTTCAAGATGATGTTTTCGGGAGAAAGATAATG	1429
QY	410	CAGTCCATGGCGCGGTGMAACGAATCCATCTGCATATGATACAGCTTGGGTGCAAGA	469	Db	1341	ACRACATGTTTAAACGTCATCGTGTTCATGTTTTCGGGAGAAACGATCATG	1400
Db	384	AGATGTATGGGTATGGCGAAGCAATCCCTCTGCAATGACACTGCTTGGGTAGCAAGG	443	QY	1430	GAAGAGCTAAGACCTTCACATCAAAATCATCCAAATGCTCTTGCCAAAACCAACGCA	1489
QY	470	ATTCCGTCAATTTGACGGCTCTGGTGACCCCAATTTCCCAACGCTTCAATFGAATCTG	529	Db	1401	GAAGAGCAAACTCTGTACCGAAAGGTATCTGAGGAATGCTCTGGAATAATGTGATGCC	1460
Db	444	ATTCCAGCAGTTGATGGCTCTGACAACCCCTCACTTTCTCGAGACGTTGAATGGATCTT	503	QY	1490	TTTGATAGTGGGCTGTCAAGAGATCTTCTGGAGAGTGGAGTATGCTATAAGTAT	1549
QY	530	AACAATCAACTCCAGATGGCTCGTGGGTGAGAGTGCATTTTCTGGCGTATCAGAGA	589	Db	1461	TTTGACAAATGGGCTTTTAAAAAGAAATATTCGGGAGAGGTAGAGTATGCACTCAAAAT	1520
Db	504	CAAAATCAGTTGAAAGATGGGCTTGGGGTGAGGATTTACTTCTTGGCATATGACAGA	563	QY	1550	CCGTGGCATAGAGTATGCGCAAGATTTGGAGGCAAGAGTTACATAGAGCAATTTTGGATCA	1609
QY	590	GTTTTAAACACTCTCGCCCTCGCTCACTCTCAAAATATGGAATAAGGGCGACATTTCAA	649	Db	1521	CCCTGGCATAGAGTATGCCAAGGTTTGGAGGCTAGAGCTATATTTGAAAACATCTATGGCCA	1580
Db	564	ATACTGGCTACACTTGCATGTATTAACCTTACCCTCTGGCGTACTGGGGAGACAA	623	QY	1610	AATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTAGAGCAACGAAAATAT	1669
QY	650	GTGCGAAAGGGTTGAGTTGTGAGAAACACATGGAAGAAATGAGACCAAGCTGAC	709	Db	1581	GATGATGTGGCTTGGAAAACTGTATATATGATGCCATACATTTCGAATGAAAAGTAT	1640
Db	624	GTACAGAAAGGTATTGAAATTTCTTCAGGACACAAGCTGGAAGATGGAAGATGAACTGAT	683	QY	1670	TTGAGCTGGCCAAATTTGGACTTCAATATGTGCGAGGCTTACACAAAGAGACTCAA	1729
QY	710	AATCAGAGCCAGTGGATTCGAGTCTGCTGCAATGTTTCTGCAATGTAGATGAAGCAAAAGC	769	Db	1641	TTAGAACTAGCGAAACTGGACTTCAATAGGTGCACTCTATACACAAACAGAGCTTCAA	1700
Db	684	AGTATAGGCCAAGTGGATTTGAATAGTATTTCTGCAATGCTAAGAGACTAAATC	743	QY	1730	CACATTTGTGAGCTGGTGAGAGAAATCGGGATTTCAATGATTTTACATTCACCCCGCAGCGG	1789
QY	770	TTGGGATTTGGATCTTCTTATCACCTCCCTTTTCACTCCCAATCCACAAAGCGCCAG	829	Db	1701	GATCTTCGAAGGTGTGGAAATCATCCGGTTTACCGGATCTGAATTTCACTCGTAGCGT	1760
Db	744	TTAGGCTTGGATCTGCTTTACGATTTTGGCAATCTCTGAACCAATCATCGAAAGCGGAG	803	QY	1790	CCTGTGGAAATGTATTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCCTGCTCT	1849
QY	830	AAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCATTAACATCAGACGGGTTGCTCTAC	889	Db	1761	GTGACGGAATAATATTTCTCACCGCATCCTTTATCTTTGAGCCGAGTTTCTTAAGTGC	1820
Db	804	GCTAAGCTTAAAGAGATTTCCCACTGATGTTCTATGCTCCCTTCCAAACACAGTTATTGTAT	863	QY	1850	AGAAATGCTATGCCAAGACTTCTTTCCTCGCAGTTATTTAGAGCATCTTTACGACACC	1909
QY	890	TCTCTGAGGTTTGCAGATGTTGGTGTGCAAGAGATGCAAAATCTTCAATCAAGA	949	Db	1821	AGAGAGTTTATACAAAACCTTCAATTTCACTGTTATTTAGATGATCTTTATGACGCC	1880
Db	864	CTTTTGGAGGTTTACAGAAATAGTAGACTGCGAGAAATATGAACTTCAATCCAAG	923	QY	1910	CACGGATCTCTGGATGATCTTAAATGTTTCTCTGAAGCGGTCCGAAGATGGGATATCTCT	1969
QY	950	GACGGATCAATTTTAAAGTCCCTGCACTCTACCTGCTTGTCTTCAATGCACTCAAAAC	1009	Db	1881	CATGATCTTTAGACGATCTTAAAGTTTTCACAGAATCAGTCAAAAGATGGGATCTATCA	1940
Db	924	GATGGATCAATTTCTCAGCTCTCGGCATCTACAGCGGCTGTATTTCTATGCGTACAGGNAAC	983	QY	1970	GTGCTGGATAGCGTTCCGGATATCAGTTGAAGTTTCTCTTAGGGCTGTACAAACACA	2029
QY	1010	AAAGGATGCTCCACTTCTCAACTCTGCTGCTCAGCAAAATTTTGGCGACTACGTTCTCTGC	1069	Db	1941	CTAGTGGAA---CCAAATGCCACAAACAAATGAATAATGTTTGTGGGTTCTACATACT	1997
Db	984	AAAAAGTCTTGGATTTCTTGAACCTTGTCTGAGAAATTCGAAACCAATGTGCTTGT	1043	QY	2030	GTGAATGGATTTTGGAAAAGATGGACTCAAGGAAACAAAGCCGCTGATGTGTGGGCTATCTT	2089
QY	1070	CATTACCCACTTGATCTATTATTGAACGCTCTGGGCTGTGATACAGTTGAAACGCTTGGGA	1129	Db	1998	TTTATGATATAGCAAAAGAGAGCGTAGAGCGCAAGGGCGGATGTGTAGGCTACATTT	2057
Db	1044	CATATCCGCTTGATCTATTATTGAACGTTTGTGGGGGCTTGATACAGTTGAGCGCTAGGT	1103	QY	2090	CGAAAAGTATGGAGGGCTTGTCTCGCATCGTATACCAAGAGAGCCGAATGGTCGGCAGCA	2149
QY	1130	ATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGATATCTGATGTGATGACTGG	1189	Db	2058	CAAAATGTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGAGAGCAATGGTCTGAAGCT	2117
Db	1104	ATCGATCGCTATTTCAGAAAGAGATCAAGGAAAGCATTTGGATATGTTTACGCGCAATGG	1163	QY	2150	AAAGTATGCCGACATTTCAACGAATATGTGGAATAATGCCAAAGTGTCCATAGCACTTGGC	2209
QY	1190	GACGCGAAGAGCGGTGGATGGGCAAGATGCAATCTTCTGATGTGATGACACT	1249	Db	2118	AAATATGTGCCATCTCTTCAATGNATACATAGAAATGCGAGTGTGTCAATAGCATTTGGGA	2177
Db	1164	GA---CGAAAGAGCAATGGATGGGCGAGAGAAATCTGTTCTTGATATTTGATATACA	1220	QY	2210	ACAGTCTGTAATAACTCAATCTTTTTCATCTGGAGAAATTAATCTCTGATTAATTTTACAG	2269
QY	1250	GCCATGGGCTTATGAATCTTGAGACTTTCATGGATACAAATGTATCTTCAGATGTTCTGGAG	1309	Db	2178	ACAGTCTTCTCATATTAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTACTCTCC	2237
Db	1221	GCCATGGGCTTTCGAATCTTGAGATTACATGGATACAAATGTATCTCTCAGATGTTTAAAA	1280	QY	2270	CAAGTATAGCTTCCGTTCCAAATTTCTGCTCTTGTCTTGTCTTGTAGCTGAGCCACTTAATCAAT	2329
QY	1310	AAATTCAGACGACGAAAGGAGACTTCTTTTGTCTTTCGGCTCAACGCAAAATTTGGTGTG	1369	Db	2238	AAAATTTGATCGGAAATCTAGATTCTTCAACTCATGGCTTAAACAGGGCGTTTGGTGAAT	2297
Db				QY	2330	GACACCAAGACTTACAGGCGCGAGAGAAACCGGTGGTGAATTTGGTTTCCACGCTACAGTGC	2389
QY				Db	2298	GACACCAAACTTATCAGGCGAGAGAGGTCAAGGTGAGGTGGCTTCTGCGACATACTGT	2357
Db				QY	2390	TACATGAGGGAATAATCCGGAGTGCACAGAGGAAGAGCTCTAAGTCAATGTTTATGTTATC	2449

```
Db 2358 TATATGAAGCACCATCTCTGAAGAAGAGCTCTACACATGCTATAGTGC 2417
QY 2450 ATGCAACACGCTGAAGGAATGAAATGGGAGTGGCCCAACCCAGGAGCAATGCCCCA 2509
Db 2418 ATGCAAAATCCCTCGAAGAGTGAATAGGAGTGTGTA-----TAACAAAATACCG 2471
QY 2510 TTGTGTGTGAAGAGCTGCTGTTCAACACTGCAAGAGTGAAGAGTGTGTTTATATGTAC 2569
Db 2472 GATATTTACAAGAGCTGGTTTGAACCTGCAAGATAATGCACTCTTTATATGCAA 2531
QY 2570 AGAGATGGCTTGGTATCTC---TGACAAAGAGATGAAGACCATGTGCACCGAATCTTT 2626
Db 2532 GGGAGTGGTTTGACACTATCACATGATATGAAATTAAGAGCATGTCAAAAATTTGCCTC 2591
QY 2627 TTGATCTCTGTGGGTAGCAT 2648
Db 2592 TTCCAACCACTTGCTAGATTA 2613

RESULT 8
AGU50768 2861 bp mRNA linear PLN 02-AUG-1996
LOCUS Abies grandis abietadiene synthase (ac22) mRNA, complete cds.
DEFINITION
ACCESSION U50768
VERSION U50768.1 GI:1477583
KEYWORDS
SOURCE Abies grandis
ORGANISM Abies grandis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
REFERENCE 1 (bases 1 to 2861)
AUTHORS Scoffer-Vogel, B., Wildung, M.R., Vogel, G. and Croteau, R.B.
TITLE Direct Submission
JOURNAL Submitted (06-MAR-1996) Mark Wildung, IBC, Washington State
University, Pullman, WA 99164-6340, USA
FEATURES
source
1..2861
/organism="Abies grandis"
/mol_type="mRNA"
/db_xref="taxon:46611"
/tissue_type="wounded stem"
1..2861
/gene="ac22"
3..2609
/gene="ac22"
/notes="diterpene cyclase"
/codon_start=1
/product="abietadiene cyclase"
/protein_id="AAB05407.1"
/db_xref="GI:1477584"
/translators="MAMPSLSSQIPTAAHLTANAQSIPIHFTLLNAGSSAKRRS
LYLRWKGSKNIIACVGGGATSPVYQSAEKNDLSSTLVKREPPPGFWKODLIDSL
TSCHKVAASDEKRIETLISIKMFRGCMGNETPSAYDTAWARIPAVDGSNDPHFP
ETWEMILONLQKDSGEGVFYLAIDRLATLACIITLTLWTGETQVOKGIEFFRTO
AKMEADASHRSRSGEIVFPAMLEKILGLDPLDPLKQIIEKREKLRIFTD
VLVALPTLLSLLEGQEIWDQIMKLOSKGDSFLSSPASTAAVFMRTGNKCLDFL
NFKLVGNFVPCPLDLRFLRWAVDVTVERLIDIRHFKBEIEKALDLYVSHWDERGI
GWARENPPVIDDTAMGLRLRLHGYNVSDVLKTRFDENGEPFCFLGOTQRGVTDML
NNVRCSHVSPGPTIMEEAKLCTERYLRNALENVDAPDKWAFKNIRGEVEYALKYVP
HKSMPLRBSYIENCPDDVWLKTVYMPYISNEKYLEKLAKLDNFVKQSIHQTELQ
DAHGSDDLKLFTEYSVKRWDLSDVQPMQMKICFVGFTNTENDIAKEGRGRGRDVL
GYTONKVMQLEAYTKEAESVLYPSENEYIENASVSLGTVLVLISALPTGEVLT
DEVLKIDRESRFLQMLGTGRVNDTKYQAEERGQGEVASAICQYMKHPKISBEBA
LOHVSVMENALBELNREFVNNKIPDIYKELVFETARIMQLFYMQDGLTLSDMEIK
EHNKCLFQPA"

Query Match 40.3%; Score 1091.4; DB 8; Length 2861;
Best Local Similarity 68.4%; Pred. No. 3e-291;
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;
QY 350 GTCGATGAGACAATCAATAAAGAATCCAGACATTTGGTGAAGGAATCCAGTGCATGTTT 409
```





Db 778 CTTGCTTCTACTGCTGCTGTTTTTATGTCACACAGGAGACGCGAAATGTCCTTGAATTTCTTG 837  
Qy 1031 AACTTCGTGCTCAGCAAAATTTGGCGACTAGCTTCCTTGGCAATTTACCACTTGTATCTATTT 1090  
Db 838 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGCTATCTCTGTGATCTGCTG 897  
Qy 1091 GAACGCTCTGGGCTGTGATACAGTTTGAACGCTTGGGAATCGATCGCTATTTCAAGAA 1150  
Db 898 GAACGCTGTGATGTAGATAATATTTAGCGCTTGGAACTTATAGACHCTTTGAAAG 957  
Qy 1151 GAAATCAAGAATCTCTCGGATTTAGTTTATPAGTACTTGGGACGCCGAAAGCGCTGGGA 1210  
Db 958 GAAATCAAGGAGCTCTCTGATTTATTTACAGGCAATTTGGAACG---AAGAGCAATTTGG 1014  
Qy 1211 TGGGCAAGATGCAATCTTATCTGATGTGATGACACTGCGCAATGGTCTTAGAATCTTG 1270  
Db 1015 TGGGCGAGATAAATCCCATAGCAGATCTTGGACCACTGCTTTGGGATTTTCGATTTGCTT 1074  
Qy 1271 AGACTTTCATGGATACAAATGATCTTCAGATGTTCTGGAGAAATTTACAGACGAGAAAGGA 1330  
Db 1075 CGGCTGCATAGGTACAAATGATCTTCAGCAATTTTGGACAACTTTCAAGATGCAATGGG 1134  
Qy 1331 GACTTCTTTTGTCTGCGGTCAACGCAAAATTTGGTGTGACCGATTAATCTTAACTTTAT 1390  
Db 1135 AAATTCATTTGCTCGACCGGTCAATTTCAACAAAGATGTAGCAAGCATGCTGATCTTTAT 1194  
Qy 1391 AGATGTTCAAGATATGTTTTCGGGAGAAAGATTAAGTGAAGCACTTAAGCTTCACT 1450  
Db 1195 AGAGCTTCCAGCTCGCATTTCCGGGAGAAACATTTCTTGATGAAGCTAAAGCTTCGCT 1254  
Qy 1451 ACAATTCATCTCCAAATGCTCTTGCACAAACCAACGCAATTTGATAAGTGGGCTGCAAG 1510  
Db 1255 ACTAATAATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAACAAACAA 1314  
Qy 1511 AAGGATCTTCTCGGAGAGTGGAGTATGCTTAAAGTATCCGTGGCATAGAAATGATGCCA 1570  
Db 1315 CAAACCTTGAGCCACAGAGATCAATACGCGCTGAGACTTCTTGGCATGCCAGTGTCCG 1374  
Qy 1571 AGATTGGAGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAG 1630  
Db 1375 AGAGTGGAGCAAGAGACTACTGTCAAGTGTATCGGCCAGATTTATGACGCGATPAGCAAA 1434  
Qy 1631 ACTGTGTATAAGTGTATATGTGACCAAGCAAAATATTTGAGCTGCGCAATTTGGAC 1690  
Db 1435 TGGTGTACAAAGTACCTACGTGGAACAATGAAAGTGTTTAGAGCTGGGAAATTAGAT 1494  
Qy 1691 TTCAATATGTTGAGCGCTTTACACCAAAAGGAGACTCAACACATTTGTACAGTGGTGAGA 1750  
Db 1495 TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGNAATGTTACCAGCTGTTTGA 1554  
Qy 1751 GAAATCGGATTTCAATGATCTTACATTCACCGCCAGCGGCTGTGAAATGTATTTCTCA 1810  
Db 1555 GATTCCGGGTGTCACATTTTACCTTTCGCTCGGAGAGCGCGCTGGAATTTACTTCTTA 1614  
Qy 1811 GTGGCGGTAGTATGTTTGGCCAGATTCGCTGCTTGTAGATTCCTATGCAAGACT 1870  
Db 1615 GTAGCGGGGAGCTATGAACCCCAAGTATGCCAAATGCAAGTTCCTCTTTACAAAGTG 1674  
Qy 1871 TCTTGCTCGCAGTTATTTAGACGATCTTTACGACACCCAGCTCTCTGATGATCTT 1930  
Db 1675 GCATGCTTCAGACTGTTCTGGACGATGTATGACACTTATGCACTTATGCACTGATGATG 1734  
Qy 1931 AAATGTTCTTGAAGCGGTCCGAAGATGGGATATCTGTGTGATGATGCGGTTCGGAT 1990  
Db 1735 AAGCTATTCACTGAGGCTGTGAGAAGATGGGACTCTCTCTTTACAGAAACCTTCCAGAC 1794  
Qy 1991 AATCAGTTGAAATTTGCTTCTTAGGGCTGTACACACAGATGATGATTTGAAAAGAT 2050  
Db 1795 TAT---ATGAACATATGTTACCAAAATCTATATGACATAGTTCCAGAGGTGGCTTGGAG 1851  
Qy 2051 GGAATCAAGGAACAAGCGCGTATGCTGGGCTATCTTCGAAAGATGATGGAGGGCTTG 2110  
Db 1852 GCAGAGAGGAACAGGGCGTGAATTTGGTCAGCTTTTTCGAAAGGATGGAGGATTTAT 1911

Qy 2111 CTGCGATCTATACCAAGACCGGAATGTTGGCGAGCAAGATATGTGCGCACTTCAAC 2170  
Db 1912 CTTCTGGGTATTTATGAAGAAGCTGAATGTTAGCTGCTGAGTATGTGCTTACCTTGGAC 1971  
Qy 2171 GAATATGTGGAATAATGCCAAAGTGTCCATAGCACTTTGCGACAGTCTGTAACCTCAATC 2230  
Db 1972 GAGTACATAAGNATGGAATCAGATCTATCGGCCAACGATATATCTTCTGTTAGTGGATG 2031  
Qy 2231 TTTTTCAC---CTGGAGAAATTTACTTCTGATTTACATTTTACAGCAAGTACCTTGGTCC 2287  
Db 2032 TTGATAATGGATGGGCAACTCTCTTTGCAAGAGGCAATTAGAGAAAGTAGATTTACAGGA 2091  
Qy 2288 AAATTTCTGC-----ATCTTGTGCTTTGACTGGAGGACTAATCAATGACACCAAGACT 2341  
Db 2092 AGACGTGTTCTCACAGAGCTGAATAGCTCATTTTCCCGCTGGCGGATGACACGAAGACA 2151  
Qy 2342 TACCAGCCGAGAGAAACCGTGGTGAATTTGGTTCAGAGCTACAGTGCCTACATGAGGAA 2401  
Db 2152 TATAAGCTGAGNAGGCTGTTGGAGAAATTTGGCTCCAGCATTTGAATGTTACATGNAAGAC 2211  
Qy 2402 AATCCGAGTGCACAGAGAAAGCTCTAAGTCAATGTTTATGTTATCATCATCGAACGCA 2461  
Db 2212 CATCTGAAATGTACAGAGAGAGGCTCTCGATCACAATCTATAGCATTTCTGGAGCCGCG 2271  
Qy 2462 CTGAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2521  
Db 2272 GTGAAGGAATCTGCAAGAGAGTTCCTGAAGCCCG---ACGACGTCCCATTTGCGCTGCAAG 2328  
Qy 2522 AGACTCTGTTCAACACTGCAAGAGTGTAGTGCAGTCTGTTTATATGTACAGAGATGGCTTT 2581  
Db 2329 AGATGCTTTTCGAGAGACAGAGTGCAGTGGTGATTTCAAGGATGAGATGATTC 2388  
Qy 2582 GGTATCTTGACAAAGAGATGAAGACCATGTGACCGCAACTCTTTTTCGATCTCTGTGGCG 2641  
Db 2389 GGTGTTTCCAAATTTAGAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCCA 2448  
Qy 2642 TAGCA-----TACTGATATATATATATATTTTCAATTCATTCACCAAAAAA 2691  
Db 2449 CTGTAATCAAAATAGTTGCAATAAATTTGAATATATCAACTATGTTTTCACAAAAA 2508  
Qy 2692 AAAAAAAAAAAAAA 2705  
Db 2509 AAAAAAAAAAAAAA 2522

RESULT 10  
BD227671  
LOCUS BD227671  
DEFINITION Synchases.  
ACCESSION BD227671  
VERSION BD227671.1 GI:33037441  
KEYWORDS JP 2002526066-A/23.  
SOURCE Abies grandis  
ORGANISM Abies grandis  
Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
REFERENCE 1 (bases 1 to 2424)  
AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.  
TITLE Synchases  
JOURNAL Patent: JP 2002526066-A 23 20-AUG-2002;  
UNIVERSITY OF KENTUCKY RESEARCH DEPARTMENT, THE SALK INSTITUTE FOR  
BIOLOGICAL STUDIES SECRETARY DEPARTMENT OF HEALTH AND HUMAN  
SERVICES  
COMMENT OS Abies grandis (giant fir)  
PN JP 2002526066-A/23  
PD 20-AUG-2002  
PF 17-SEP-1999 JP 2000574228  
PR 18-SEP-1998 US 60/100993,22-APR-1999 US 60/130628 PR  
23-AUG-1999 US 60/150262  
PI JOSEPH CHAPPELL,KATHLEEN R MANNA,JOSEPH P NOEL,COURTNEY M PI  
STARKS  
PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N5/



```
QY 2342 TACCAGCCGAGAGAACCGTGGTGAATTGGTTTCCAGCGTACAGTGCTACATGAGGAA 2401
Db 2048 TATAAAGCTGAGAAAGGCTCGTGAGAAATTGGCGTCCAGCAATTGAATGTTACATGAAAGAC 2107
QY 2402 AATCCCGAGTGCCAGAGGAAGAGCTCTAAGTCAATGTTTATGGTATCATCGACAACGCA 2461
Db 2108 CATCTGTAATGTACAGAGGAGAGGCTCTCGATCACATCTATAGCATTCTGAGCCGGCG 2167
QY 2462 CTGAAGGAATTGAATTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGTGAGA 2521
Db 2168 GTGAAGGAACCTGCAAGAGAGGTTTCTGAAGCCCG---ACGACGTCCCATTTGCGCTGCAAG 2224
QY 2522 AGACTGCTGTCAACACTGCMAGAGNGATGCGAGCTGTTTATATGTATACAGAGATGCTTT 2581
Db 2225 AGATGCTTTTCGAGGAGCAAGAGTGACGATGGTGATATTCAGGATGGAGATGGAATTC 2284
QY 2582 GGTATCTCTGACAAAGAGATGAAAGACCATGTGAGCCGAACCTTTTTCGATCCTGTGGCG 2641
Db 2285 GGTGTTCCAAATTAGAGTCAAGATCATATCAAGAGATGTCTCATTGNAACGCTGCCA 2344
QY 2642 TAGCA-----TACTGATATATATATAATATATATATATATATATATATATATATAT 2688
Db 2345 CTGTAATCAAAATAGTGGCAATAATAATTGAATAATGTCAACTATGTATTTTCACAAAAAA 2404
QY 2689 AAAAAAAAAAAAAA 2705
Db 2405 AAAAAAAAAAAAAA 2421

RESULT 11
LOCUS AR222107 2424 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 15 from patent US 6429014.
ACCESSION AR222107
VERSION AR222107.1 GI:23329481
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2424)
AUTHORS Steele, C.L., Bohlmann, J. and Croteau, R.B.
TITLE Monoterpene synthases from grand fir' (Abies grandis)
JOURNAL Patent: US 6429014-A 15 06-AUG-2002;
FEATURES
source Location/Qualifiers
1..2424
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 27.0%; Score 730.2; DB 6; Length 2424;
Best Local Similarity 59.6%; Pred. No. 5.6e-191;
Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;

QY 434 AATCCATCTGCATATATACAGTCTGGTGGCAAGAAATCCGTCAATTGACGGCTCTGGT 493
Db 134 ACTCCATCTGCTTATGACACAGATGGGTAGCGAGGGTGCCGCCATTGATGGCTCTGCT 193
QY 494 GCACCCCAATTTCCCAAAACGCTTCAATGGATTCTGAACAATCAACTGCCAGATGCTCG 553
Db 194 CGCCCGCAATTTCCCAAAACAGTTGACTGGATTGGAATAACAGTTAAAGATGTTCA 253
QY 554 TGGGGTGAGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACAACCTCTGCGCTGCTTC 613
Db 254 TGGGGAATTCAGTCCCACTTTCTGCTGCTCGACCGTCTTCTTGCCACTCTTTCTTGTT 313
QY 614 CTCACCTCTCAAAATATGGAATAGGCGACATTCAGTGCGAGAAAGGGTTGATTTG 673
Db 314 CTGTGCTCTTAAATTTGGAACGTTGGGGATCTCNAAGTAGAGAGAGGAAATTGAATTCATA 373
QY 674 AGAAAAACATGGAAGAAATGAAGGACGAGCTGCAATCACAG---GCCAAGTGATTC 730
Db 374 AAGAGCAATCTGGAACTAGTAAAGGATGAACCGATCAAGATAGCTTGGTAACAGACTTT 433
```

```
QY 731 GAGTCGTCTTCCCTGCAANTGTAGATGAAGCAAAAAGCTTGGATTGGATCTTCCCTTAT 790
Db 434 GAGATCATATTTCCCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGAGCTTCCCTAC 493
QY 791 CACCTCCCTTTTCAATCTCCCAATCCACCAAAAGCGCAGAAAAGCTTCAAAAAGATTCCC 850
Db 494 GACCTGCCCTTATATACATCTGTTGCAGACTAAACGGCAGGAAAGATTAGCAAACTTCA 553
QY 851 CTCATATGTTCTTCAATCAACCATACAGCGGGTGTCTCTACTCTCTGAGGGTTTGCAGAT 910
Db 554 AGGGAGAAAATTTATGCGGTTCGTCGCAATTGTTGTATTTCTTTAGAGGGAATACAGAT 613
QY 911 GTGGTGAGCTGCGAAGAGATCAAAATCTTCAATCAAGAGACGATCATTTTAAAGCTCC 970
Db 614 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAGTCAGATGGGTCTTTCTTAAAGCTCA 673
QY 971 CTTGCATCTACTGCTTGTGTCTTTCATGCACTCAAAAACAAACGATGCTCCCTTCTC 1030
Db 674 CTTGCTTCTACTGCTGCGTTCATGCAACAGGAGACGCGAAATGCTTGAATCTTG 733
QY 1031 AACTTCGTCTCAGCAAAATTTGGCGACTAGTTCCTTGGCAATTAACCACTTGTATTT 1090
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTCCCTGCTGATCTCTGTGATCTGCTG 793
QY 1091 GAACGCTCTGGGCTGTCGATACAGTTGAACGCTTGGGATCGATCGTATTTCAAGAAA 1150
Db 794 GAAAGCTTGTGATCGTAGATAATATTGTACGCTTGGAAATCTTATAGACATTTGAAAG 853
QY 1151 GAAATCAAAAGAAATCTCTGATTACGTTTATAGTACTTGGGACGCCCAAGAGCGTGGGA 1210
Db 854 GAAATCAAGAAAGCTCTTGTATGTTTACAGGCATTGNAACG---AAAGAGAAATTTGGG 910
QY 1211 TGGGCAAGATGCAATCTTCTGATGCGATGACACTGCCATGGTCTTTAGAACTCTG 1270
Db 911 TGGGGCAGACTAAATCCCATAGCAGATCTTGAGACCCTGCTTTGGGATTTTCGATTGCTT 970
QY 1271 AGACTTCATGGATACAAATGATCTTTCAGATGTTCTGGAGAAATTTCAAGACGAGAAAGGA 1330
Db 971 CGGCTGATAGGTACAAATGATCTCTCAGGCCATTTTGTGACAACTTTCAAGATGCGAATGGG 1030
QY 1331 GACTCTTTTGTCTTTCGCGGTCAACGCAAAATTTGGTGTGACCGATTAATCTTAAACCTTTAT 1390
Db 1031 AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAGCATGCTGAATCTTTAT 1090
QY 1391 AGATGTTCAAGATATGTTTTCGGGAGAAAAGATAATGGAAGAGCTTAAGACTTCACCT 1450
Db 1091 AGAGCTTCCAGCTCGCAATTTCCGGAGAAAAACATTTCTTGATGAAGCTTAAAGCTTCGCT 1150
QY 1451 ACNAATCATCTCCAAATGCTCTTTCACAAAACAAACGCAATTTGATAGTGGGCTGTCAAG 1510
Db 1151 ACTAAATATTTGAGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGCAATGGAACAACAA 1210
QY 1511 AAGGATCTTCTCGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAGTATGCCA 1570
Db 1211 CAAACCTGAGCCCAAGATCAAAATAGCGCTGAGAACTTCTTGGCATGCGAGTGTTCG 1270
QY 1571 AGATTGGAGCGCAAGAGTTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG 1630
Db 1271 AAGTGGAGCAAGAGATACCTGCAAGTGTATCGCCAGATTTATGCAACCATAGCAAAA 1330
QY 1631 ACTGTGTATAAGATGCTATATGTAGACCAAGAAAATATTTGGAGCTGGCCAAATTTGGAC 1690
Db 1331 TGGTTTTACAAAGCTACCCCTACGTTGAACAATGAAAGTTTTTTAGAGCTGGGAAAAATTAGAT 1390
QY 1691 TTCAATATGCTGAGCGCTTTACACCAAAAGGAGACTCAACACATTTGTACGCTGTGTGAGA 1750
Db 1391 TTCAACATTTACAGTCCATCCCAAGAGAAATGAAGATGTTTACCAGCTGTTTGA 1450
QY 1751 GAATCGGGATTTCAATGATCTTCAATTCACCGCCAGCGGCTGTGAAATTTGATTTCTCA 1810
Db 1451 GATTCCGGGTGCGCACTATTTCACCTTCGCTCGGAGAGCGCGCTGGAAATTTCTACTCTTA 1510
QY 1811 GTGGCGGTAGTATGTTTGGAGCAGAAATTCGCTGCTTGTAGAAATTTGCTTATGCCAAGACT 1870
```





Db 434 GAGATCATATTTCTTCTCTGTATAGAGAAGCTCAATCTCTGCGCTCGGACTTCCCTAC 493  
 Qy 791 CACCTCCCTTTTCATCTCCCAATCCACAAAAGCCGAGAAAAGCTTCAAAAGATTCCC 850  
 Db 494 GACCTGCTTATATACATCTGTGAGACTAAACCGCAGGAAAGATTAGCAAACTTTCA 553  
 Qy 851 CTCAATGTTCTTCAATACCATCAGACGGGTGCTCTACTCTCTGAGAGGTTTCAAGAT 910  
 Db 554 AGGAGGAGAAATTTATGCGGTTCGGTCCGTCATTTGTTGTTATCTTTAGAGGGAATACAAGAT 613  
 Qy 911 GTGGTGAGCTGGCAGAGATCAAAATCTTCAATCAAGAGAGGATCATTTTAAGTCCC 970  
 Db 614 ATAGTTGAATGGGAACGAATATATGAAGTTTCAAGTTCAGGATGGTCTTTCTTAAAGTCA 673  
 Qy 971 CTTGATCTTACTGCTGCTGCTTCTTCAATGACACTCAAAACAAACAGATGCTCCACATTTCTC 1030  
 Db 674 CTTGCTTCTTACTGCTGCTGCTTCTTCAATGACACTCAAAACAAACAGATGCTCCACATTTCTC 733  
 Qy 1031 AACTTCTGCTCAGCAAAATTTGGCGACTACGTTCTTGCATTTACCCACTTGATCTATTT 1090  
 Db 734 AACAGTGTGATCAAGTTTGGAAATTTTGTCTGCTGCTGTATCTCTGCTGATCTGCTG 793  
 Qy 1091 GAACGCTCTGGCTGTGATACAGTTCGATCAGTTCGATCAGTTCGATCAGTTCGATCAGTTC 1150  
 Db 794 GAACGCTCTGGCTGTGATACAGTTCGATCAGTTCGATCAGTTCGATCAGTTCGATCAGTTC 853  
 Qy 1151 GAAATCAAGAATCTCTGATTAAGTTTATAGTACTGAGGAGCGGAAAGCGCTGGGA 1210  
 Db 854 GAAATCAAGAATCTCTGATTAAGTTTATAGTACTGAGGAGCGGAAAGCGCTGGGA 910  
 Qy 1211 TGGGCAAGATCAATCTTATCTGATGCTGATGACACTGCTGATGACACTGCTGATGACACT 1270  
 Db 911 TGGGCAAGATCAATCTTATCTGATGCTGATGACACTGCTGATGACACTGCTGATGACACT 970  
 Qy 1271 AGACTTCTGATCAATCTTATCTGATGCTGATGACACTGCTGATGACACTGCTGATGACACT 1330  
 Db 971 CGGCTGATGATCAATCTTATCTGATGCTGATGACACTGCTGATGACACTGCTGATGACACT 1030  
 Qy 1331 GACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1390  
 Db 1031 AATTTCAATTTGCTCGACCGCTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT 1090  
 Qy 1391 AGATGTTCAAGTATGTTTTCGGGAGAAAGATTAATGGAAGAGCTTAAGACTTCACT 1450  
 Db 1091 AGAGCTTCCAGCTGCTGCTTCCCGGAGAAACATTTCTTGATGAAGCTTAAAGCTTCGCT 1150  
 Qy 1451 ACATATCATCTCCAAATCTCTTCCCAAAACAAAGCATTTGATGAAGTGGCTGTCAAG 1510  
 Db 1151 ACTAATATTTGAGAGAGCTCTTGAGAAAGTGAAGTTCAGTGCATGGAAACAA 1210  
 Qy 1511 AAGGATCTTCTGAGAGGTGAGTATGCTATAAAGTATCCGTTGGCATAGAAATGCGCA 1570  
 Db 1211 CAACCTTCCAGCAAGAGATCAATACGCTGAGACTTCTTGGCATGCCAGTGTTCG 1270  
 Qy 1571 AGATGGAGGCAAGATTAATGAGCAATTTGGATCAATGATGCTGCTGCTGGGAG 1630  
 Db 1271 AGAGTGGAGCAAGAGATCACTGTCAAGTGTATCGCCAGATTTATGACGCTAGCAAAA 1330  
 Qy 1631 ACTGTGATTAAGATGCTATATGAGCAACGAAATATTTGAGCTGCGCAATTTGGAC 1690  
 Db 1331 TCGGTTTACAGTACCTTACGTAACATGAAGTTTATGAGCTGGGAAATTTAGAT 1390  
 Qy 1691 TTCAATATGTTGAGGCTTTACACCAAAAGGAGACTCAACACATTTGCTAGCTGGTGAGA 1750  
 Db 1391 TTCAATATTTCCAGTCCATCCACCAAGAGAAATGAAGAAATGTTACAGCTGGTTTGA 1450  
 Qy 1751 GAATCGGATTTCAATGATCTTACATTTCAACCGCAGGCTGCTGGAATGTTTCTCA 1810  
 Db 1451 GATTCGGGTTGCCACTATTTCACTTCTGCTGCGGAGAGCGCTGGAAATTTCTACTTTCTA 1510  
 Qy 1811 GTGGCGGTTAGTATGTTTGGAGCAGAAATTCGCTGCTTGTAGAAATTCGCTATGCAAGACT 1870  
 Db 1511 GTAGCGCGGGAGCTTATGACCCAGTATGCCAATGCAAGTTCCTCTTTACAAAGTG 1570

Qy 1871 TCTTCCCTCGCAGTTATTTCTAGACGATCTTTACGACACCCACGAGATCTCTGAGATGATCTT 1930  
 Db 1571 GCATGCTTGCAGACTGTTCTGACGATATGTATGACACTTATGGAACCCCTAGATGAATTG 1630  
 Qy 1931 AAATTTGTTCTCTGAACCGGTCCGAAGATGGGATATCTCTGCTGCTGATAGCTTCCGGAT 1990  
 Db 1631 AAGCTATTTCACTGAGGCTGTGAGAAGATGGGACCTCTCTCTTTACAGAAACCTTCCAGAC 1690  
 Qy 1991 AATCAGTTGAAAGTTTGTCTCTAGGCTGTACAAACACAGTGAATGGAATTTGGAAGAAT 2050  
 Db 1691 TAT---ATGAAGTATGTTACCAATCTATATGACATAGTTTACAGAGTGGCTTGGAG 1747  
 Qy 2051 GGAATCAAGGAAACAGGCGGTGATGCTGGGCTATCTTTCGAAAAGTATGAGAGGCTTG 2110  
 Db 1748 GCAGAGAAGGAACAGGCGGTGAAATTTGCTGAGCTTTTTCAGAAAGGATGGGAGGATTA 1807  
 Qy 2111 CTCGATCTGATACCAAGAGGCGGATGCTGGGAGCAAGATGATGCTGCGGACATTTCAAC 2170  
 Db 1808 CTTCTGGGTTATATGAAGAGCTGAATGTTAGCTGCTGAGTATGCTGCTACCTTGGAC 1867  
 Qy 2171 GAATATGTGGAATAATGCCAAAGTGTCCATAGCACTTGGCAGACTGCTACTAACTCAATC 2230  
 Db 1868 GAGTACATAAAGATGGAATCACAATCTATCGCCCAAGTATATCTTCTGTTGAGTGGAGTG 1927  
 Qy 2231 TTTTTC---CTGGAGAAATTAATCTTCTGATTAATTTTACGCAAGTACCTTCCGTTCC 2287  
 Db 1928 TTGATAATGAGTGGGCAACTCTCTTTCGCAAGAGGCAATTTAGAGAAAGTAGATTTATCCAGGA 1987  
 Qy 2288 AATTTCTGC-----ATCTTGTCTTTGACTGGACGACTAATCAATGACACCAAGACT 2341  
 Db 1988 AGAGTGTCTTACAGAGCTGAATAGCTCATTTCCGCTGCGGATGACAGCAAGACA 2047  
 Qy 2342 TACCAGGCGGAGAGAAACCGTGGTGAATTTGGTTCAGCGTACAGTGTACATGAGGGAA 2401  
 Db 2048 TATAAAGCTGAGAGGCTGCTGGAGAAATTTGGCGTCCAGCAATTTGATGTAAGAAAGAC 2107  
 Qy 2402 AATCCGAGTGCAGAGAGAAAGCTCTAAGTCAATGTTTATGTTATCATCGCAAGCA 2461  
 Db 2108 CATCTGATGTACAGAGAGAGGCTCTCGATCATCTATAGCAATTTGAGAGCCGCG 2167  
 Qy 2462 CTGAAGGAAATCAATTTGGAGTTGGCAACCCAGCGAGCAATGCTCCCATTTGTGTGAGA 2521  
 Db 2168 GTGAAGAACTGACAGAGAGTTTCTGAAGCCG---ACGAGTCCCATTTCCGCTCAAG 2224  
 Qy 2522 AGACTGCTTTCAACACATGCAAGAGTGTGAGCTGTTTTATATGATGACAGAGATGGCTTT 2581  
 Db 2225 AAGATGCTTTTCGAGGAGACAAAGAGTGCAGTGTGATATTTCAAGGATGGAGATGATTC 2284  
 Qy 2582 GGTATCTGCAAAAGAGATGAAGACCATGTCAGCCGAACTTTTTCGATCCTGTGGCG 2641  
 Db 2285 GGTGTTTCCAAATTTAGAGTCAAGATCATATCAAAAGAGTGTCTCATTTGAACCGCTGCCA 2344  
 Qy 2642 TAGCA-----TACTGATATATATATATATATTTCAATTTCAATTTCAATTTCAATTTCA 2688  
 Db 2345 CTGTAATCAAAATAGTTGCAATAATTAATGAAATATGTCACACTATGTTTTCACAAAAA 2404  
 Qy 2689 AAAAAAATAAAAAA 2705  
 Db 2405 AAAAAAATAAAAAA 2421

RESULT 14  
 AR266981  
 LOCUS 2424 bp DNA linear PAT 10-APR-2003  
 DEFINITION Sequence 45 from patent US 6495354.  
 ACCESSION AR266981  
 VERSION AR266981.1 GI:29696436  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 2424)







```
Db 1571 GCATGCTTGACAGCTGTTCTGGACGATATGTATGACACTTATGGAACCCCTAGATGAATTG 1630
Qy 1931 AAATTGTTCTCTGAAGCGTCCGAGATGGGATATCTCTGTGCTGGATAGCGTTCGGGAT 1990
Db 1631 AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCCTTTACAGAAAACCTTCCAGAC 1690
Qy 1991 AATCAGTTGAAAGTTTGCTTCCTAGGCTGTACACACAGTGAATGGAATTTGGAAAAGAT 2050
Db 1691 TAT---ATGAAACTATGTTACCAATCTATATATGACATAGTTCACGAGGTGGCTTGGGAG 1747
Qy 2051 GGACTCAAGGAAACAAGGCGTGTATGCTGGGCTATCTTCGAAAAGTATGGGAGGCTTG 2110
Db 1748 GCAGAGAAGGAACACGGGCGTGAATTGGTCAGCTTTTCAGAAAAGGATGGGAGATTAT 1807
Qy 2111 CTCGCATCGTATACCAAGAAGCCGAATGGTCGGCAGCAAGATATGTGCCGACATTC AAC 2170
Db 1808 CTTCTGGGTTATATATGAAGAAGCTGAATGGTTAGCTGTGCTAGTATGTGCTACCTTGGAC 1867
Qy 2171 GAATATGTGGAATAAGCCAAAGTGTCCATAGCACCTTGGACAGTCTGTAACCTCAATC 2230
Db 1868 GAGTACATAAAGATGGAAATCACATCTATCGGCCAACGTATATCTTCTGTGTGAGTGAGTG 1927
Qy 2231 TTTTTC---CTGAGAAATTTACTCTCTGATTACATTTTACAGCAAGTAGACCTTCGGTCC 2287
Db 1928 TTGATAATGGATGGCAACTCTTTTGGCAGAGGCAATTAGAGAAAGTAGATTATCCAGGA 1987
Qy 2288 AAATTTCTGC-----ATCTGTGTCTTTGACTGGAAGCACTAATCAATGACACCAAGACT 2341
Db 1988 AGACGTGTTCTCACAGAGCTGAATAGCCTCAATTCGCCGCTGGCGATGACACGGAAGCA 2047
Qy 2342 TACAGGCGGAGAGAAACGGTGGTGAATTTGTTTCCAGGTCAGTGTCTACATGAGGGA 2401
Db 2048 TATAAAGCTGAAGCGCTCGTGAGAAATTTGGCGTCCAGCAATTTACATGAAGAC 2107
Qy 2402 AATCCGAGTGCACAGAGAGAAAGCTCTAAGTCAATGTTTATGTTATCATCGACACACGCA 2461
Db 2108 CATCTGATGTACAGAGAGAGGCTCTCGATCACTATAGCATTTCTGAGCGCGCG 2167
Qy 2462 CTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCAATTTGTGTGAGA 2521
Db 2168 GTGAAGAACTGACAAGAGAGTTTCTGAAGCCG---ACGACGTCCCATTCGCCCTGCAAG 2224
Qy 2522 AGACTGCTGTTCAACACTGCAAGAGTGTGACAGCTGTTTATATGTACAGAGATGCTTT 2581
Db 2225 AAGATGCTTTTCGAGGAGACAGAGTGACGATGGTGATATTCAGGATGGAGATGGAATTC 2284
Qy 2582 GGTATCTCTGACAAAGAGATGAAGACCATGTGACGCCGAATCTTTTTCGATCCTGTGGCG 2641
Db 2285 GGTGTTCCAAATTAAGAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGTGCCA 2344
Qy 2642 TAGCA-----TACTGATATTTATATATATATTTCAATTTCAATCCAAAAAAA 2688
Db 2345 CTGTAATCAAAATAGTTCGAATAATAATGAAATAATGTCAACTATGTTTTCACAAAAAA 2404
Qy 2689 AAAAAAAAAAAAAA 2705
Db 2405 AAAAAAAAAAAAAA 2421
```

Search completed: August 24, 2004, 20:44:24  
Job time : 10506.1 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 11:17:28 ; Search time 1009.04 Seconds  
(without alignments)  
11388.394 Million cell updates/sec

Title: US-10-041-007-1  
Perfect score: 2705  
Sequence: 1 atatttgcataacctgcac.....aaaaaaaaaaaaaaaaaaaaa 2705

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : 1: Geneseqn\_29Jan04:\*  
2: Geneseqn1980s:\*  
3: Geneseqn1990s:\*  
4: Geneseqn2000s:\*  
5: Geneseqn2001as:\*  
6: Geneseqn2001bs:\*  
7: Geneseqn2002as:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1091.4	40.3	2861	3 AAA38937	Aaa38937 Grand fir
2	731.6	27.0	2525	2 AAX87530	Aax87530 E-alpha-b
3	730.2	27.0	2424	2 AAX08654	Aax08654 Grand fir
4	730.2	27.0	2424	3 AAA38932	Aaa38932 Grand fir
5	730.2	27.0	2424	4 AAF73382	Aaf73382 Grand fir
6	730.2	27.0	2528	2 AAX87529	Aax87529 Grand fir
7	730.2	27.0	2571	2 AAX87531	Aax87531 E-alpha-b
8	730.2	27.0	2700	2 AAT97447	Aat97447 Pacific Y
9	642.2	23.7	2700	3 AAA38931	Aaa38931 Yew taxad
10	642.2	23.7	2700	3 AAA38931	Aaa38931 Yew taxad
11	450.4	16.7	2186	4 AAF73412	Aaf73412 Grand fir
12	416	15.4	1890	4 AAF73421	Aaf73421 Grand fir
13	412.2	15.2	2205	2 AAX08663	Aax08663 Grand fir
14	412.2	15.2	2205	4 AAF73391	Aaf73391 Grand fir
15	411.2	15.2	1865	3 AAA38933	Aaa38933 Grand fir
16	411.2	15.2	1885	2 AAX87532	Aax87532 Delta-sel
17	411.2	15.2	1885	2 AAX87533	Aax87533 Delta-sel
18	411.2	15.2	1888	2 AAX87505	Aax87505 Grand fir
19	411.2	15.2	1888	2 AAX87505	Aax87505 Grand fir
20	410.6	15.2	2196	2 AAX08643	Aax08643 Myrcene s
21	410.6	15.2	2196	3 AAA38927	Aaa38927 Grand fir
22	410.6	15.2	2196	4 AAF73371	Aaf73371 Grand fir
23	409.4	15.1	2460	3 AAZ99415	Aaz99415 Spybean c

24	409.4	15.1	2460	6 AAD40203	Aad40203 Soybean c
25	407.2	15.1	2089	2 AAX08645	Aax08645 Limonene
26	407.2	15.1	2089	3 AAA38938	Aaa38938 Grand fir
27	407.2	15.1	2089	4 AAF73373	Aaf73373 Grand fir
28	404.4	15.0	1967	2 AAX08655	Aax08655 Grand fir
29	404.4	15.0	1967	2 AAX08655	Aax08655 Grand fir
30	404.4	15.0	1967	2 AAF73383	Aaf73383 Grand fir
31	401.8	14.9	2018	2 AAX08644	Aax08644 Pinene sy
32	401.8	14.9	2018	3 AAA38922	Aaa38922 Grand fir
33	401.8	14.9	2018	4 AAF73372	Aaf73372 Grand fir
34	390.6	14.4	2429	4 AAF73413	Aaf73413 Grand fir
35	376.4	13.9	2587	2 AAT15761	Aat15761 Gibberell
36	375	13.9	1513	3 AAA69551	Aaa69551 Pinus rad
37	373.4	13.8	1634	3 AAX69644	Aax69644 Pinus rad
38	372.8	13.8	2013	4 AAF73411	Aaf73411 Grand fir
39	347.8	12.9	1785	2 AAX87536	Aax87536 Gamma-hum
40	347.8	12.9	1785	2 AAX87535	Aax87535 Gamma-hum
41	347.8	12.9	1785	3 AAA38934	Aaa38934 Grand fir
42	347.8	12.9	1977	2 AAX87506	Aax87506 Grand fir
43	346.2	12.8	1785	2 AAX87537	Aax87537 Gamma-hum
44	329.6	12.2	2784	2 AAT09940	Aat09940 Anther ea
45	306	11.3	2109	3 AAZ99417	Aaz99417 Wheat cop

## ALIGNMENTS

RESULT 1  
ID AAA38937 standard; DNA; 2861 BP.  
XX  
AC AAA38937;  
XX

25-AUG-2000 (first entry)

Grand fir abietadiene synthase DNA sequence SEQ ID NO:55.

Synthase; protein co-ordinate data; active site; modification; terpenoid;  
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
isoprenoid; breeding programme; fragrance; flavour; pheromone;  
defensive agent; pigment; antitumour; steroid hormone;  
signal transduction pathway; bile acid; affinity purification;  
photoreceptor; enzymatic synthesis; nutrient supplement;  
immunological reagent; ds.

Abies grandis.

WO200017327-A2.

30-MAR-2000.

17-SEP-1999; 99WO-US021419.

18-SEP-1998; 98US-0100993P.

22-APR-1999; 99US-0130528P.

23-AUG-1999; 99US-0150262P.

(KENT ) UNIV KENTUCKY RES DEPT.

(SALK ) SALK INST BIOLOGICAL STUDIES.

Chappell J, Manna KR, Noel JP, Starks CW;

WPI; 2000-292839/25.

P-PSDB; AAY90858.

Novel terpene synthase enzymes, useful for producing terpene

hydrocarbons, e.g. fragrances or antitumor agents, are derived from known

enzymes by specific amino acid alterations.

Disclosure; Page 439-443; 450pp; English.

The present invention describes an isolated terpene synthase (1)

comprising a region with at least 20% identity to region 285-535 of a 548

CC	amino acid (aa) sequence (1a), given in AAY90831. (1) contains nine alpha
CC	-carbon atoms (alphaC) that have interatomic distances, between each
CC	other, within tabulated ranges, have a centre point (within a sphere of
CC	radius 2.3 Angstrom) within tabulated ranges, and have an ordered
CC	arrangement of R groups (defining aa side chains), excluding specific
CC	tabulated arrangements (tables given in the specification). (1), and
CC	related enzymes, are used to produce a wide range of terpenoids (e.g.
CC	cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,
CC	flavours, pheromones, defensive agents, pigments, antitumour agents,
CC	components of signal transduction pathways, precursors of steroid
CC	hormones and bile acids, as photoreceptors and as co-factor side chains.
CC	Some syntheses with little or no catalytic activity (and nucleic acids
CC	encoding them) are used as controls in the analysis of products formed by
CC	enzymatic synthesis; as nutrient supplements; for affinity purification
CC	of isoprenoids; or to develop immunological reagents or nucleic acids for
CC	monitoring expression of terpene synthase or inheritance of the gene in
CC	plant breeding programs. The new syntheses may produce novel terpene
CC	products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent
CC	sequences used in the exemplification of the present invention
XX	
SQ	Sequence 2861 BP; 890 A; 514 C; 639 G; 818 T; 0 U; 0 Other;
Query Match 40.3%; Score 1091.4; DB 3; Length 2861;	
Best Local Similarity 68.4%; Pred. No. 1.4e-284;	
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;	
QY	350 GTGATGAGACAAATCAATAAAGATCCAGACATTTGGTGAAGGAATCCAGTGCATGTTT 409
DB	
QY	324 GTTGCAGCATCAGACGAGAGCGTATCGAGACATTAATATCCGAGATTAAGAAATATGTTT 383
DB	
QY	410 CAGTCCATGGGCGACGGTGAACGAATCCATCTGCATATGATACAGCTTTGGGTGGCAAGA 469
DB	
QY	384 AGATGATGGGTATGCGCAACGAATCCCTCTGCATATGACACTGCTTGGGTAGCAAGG 443
DB	
QY	470 ATTCCGTCAAATGACGGCTCTGGTGCACCCCAATTTCCCAACGCTTCAATGGATTTCTG 529
DB	
QY	444 ATTCCAGCAGTTGATGGCTCTGACACCCCTCACTTTCTCGAGACGGTTGAATGGATTTCT 503
DB	
QY	530 AACATCACTGCCAGATGGCTGGGTGGGTGAGGATGATTTTCTGGCGTATGACAGA 589
DB	
QY	504 CAAATACGTTGAAGATGGTCTTGGGGTGAAGGATTTCTCTTTCTGGCATATGACAGA 563
DB	
QY	590 GTTTTAAACACTCTCGCCCTCCCTCACTCTCAAAATATGGAAATGAAGCGCGCATTCAA 649
DB	
QY	564 ATACTGGCTACACTTCGATGATTAATTAACCTTACCTCTGGCGTACTGGGAGACACAA 623
DB	
QY	650 GTGCAGAAAGGGTTGAGTTGTGAGAAACACATCGAAGAAATGAAGCAAGCTGAC 709
DB	
QY	624 GTACAGAAAGGTATTTGAATTTCTTCAGGACACAAGCTGGAAAGATGAAGCTGAT 683
DB	
QY	710 AATCAGAGCCAAAGTGGATTCGAGGTCTGTCTTCTGCAATGTTAGATGAAGCAAAAAGC 769
DB	
QY	684 AGTCATAGGCCAAGTGGATTTGAATAGTATTTCTGCAATGCTAAAGGAAGCTAAATC 743
DB	
QY	770 TTGGGATTTGGATCTTCTTATACCTCCCTTTCACTCCCAATCCACCAAGCGCCAG 829
DB	
QY	744 TTAGGCTTGGATCTGCTTACGATTTGCCATTTCCCTGAAAACAAATCATCGAAAACGGGAG 803
DB	
QY	830 AAAAGCTTCAAAGATTTCCCTCAATGTTCTTCAATCAACCATCAGAGCGGTGCTCTAC 889
DB	
QY	804 GCTAAGCTTAAAGGATTTCCCACTGATGTTCTCTATGCCCTTCAACACGTTATGTAT 863
DB	
QY	890 TCTCTGGAGGTTTGAAGATGTGGTGAAGTGGCAAGAGAGATCAAAATCTTCAATCAAGA 949
DB	
QY	864 TCTTTGGAAGTTTACAGAAATAGTAGACTGGCAGAAATATGAATGAACCTTCAATCAAG 923
DB	
QY	950 GACGGATCAATTTTAAAGTCCCTGCAATCTACTGCTTGTCTTTCATGCACTCAAAAC 1009
DB	
QY	924 GATGGATCAATTTCTCACTCTCCGCACTCTACAGCGCTGTATTTATGTCGTACAGGAAAC 983
DB	
QY	1010 AAAGGATGCTCCACTTTCTCACTCTCGTCTCAGCAAAATTTGGCGACTACGTTCTCTTGC 1069
DB	
QY	984 AAAAGTGTCTGGATTTCTTGAATTTCTTGTCTTGAGAAATTCGGAACCATGTGCTTGT 1043
DB	

QY	1070 CATTACCCACTTGATCTATTTTGAACGCTCTGGGCTGTCGATACAGTTGAACGCTTGGGA 1129
DB	
QY	1044 CACTATCCGCTTGATCTATTTTGAACGCTTGTGGGGTGTGATACAGTTGAGCGGCTAGT 1103
DB	
QY	1130 ATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTTAGCTTTATAGGTACTGG 1189
DB	
QY	1104 ATCGATCGCTATTTCAAGAGAGGAGATCAAGGAAGCATTTGGATTTATTTTACAGCCATTTGG 1163
DB	
QY	1190 GACGCCGAAAGGCGTGGGATGGGCAAGATGCAATCTATTCCGTGATGTCGATGACACT 1249
DB	
QY	1164 GA---CGAAGAGGCAATTTGGATGGCGGAGAGAGAAATCTCTTCCGTGATTTGATGATACA 1220
DB	
QY	1250 GCCATGGCTTTAGAAATCCTGAGACTTCTCATGATACAAATGATATCTTCAGATTTCTGGAG 1309
DB	
QY	1221 GCCATGGCCCTTCGAATCTTCGATTTACATGATCAATGATCTCTCAGATGTTTATAAA 1280
DB	
QY	1310 AATTTACAGACGAGAAAGGAGACTTTCTTTTCTGCTTGGGTCGCAACGCAAAATTTGGTGG 1369
DB	
QY	1281 ACATTTAGAGATGAGAAATGGGGAGTTCTTTTCTTGGGTCGCAACACAGAGAGAGTT 1340
DB	
QY	1370 ACCGATAATCTTAACCTTTTATAGATGTTTCAAGATGTTTTCGGGGAGAAAGATAATG 1429
DB	
QY	1341 ACAGACATGTTAAACGTTCAATCGTTTTCACATGTTTTCATTTCCGGGAGAAACGATCATG 1400
DB	
QY	1430 GAAGAAGCTTAAGACCTTCTCACTACAAATCATCTCCAAATGCTCTTCCGCAAAACCAACGCA 1489
DB	
QY	1401 GAAGAAGCAAAACTCTGTACCGAAAGGATCTGTAGGAATGCTCTGGAAATTTGGATGCC 1460
DB	
QY	1490 TTTGATAGTGGCTCTCAAGAGGATCTTCTGAGAGGTCGATGATCTCTATAAGTAT 1549
DB	
QY	1461 TTTGCAAAATGGGCTTTTAAAGAAATATTTCCGGGAGAGGTAGATGATGCACTCAATAT 1520
DB	
QY	1550 CCGTGCCTAGAAATGATGCCAAGATTTGGAGGCAAGAGTTACATAGAGCAATTTGGATCA 1609
DB	
QY	1521 CCTGCGCATAGAGTATGCCAAGTTTGGAGGCTAGAGCTATATTTGAAACATATGGGCCA 1580
DB	
QY	1610 AATGATGCTGCTGGGGAAGAAGCTGTGTATTAAGATGCTATATGTGAGCAACGAAATAT 1669
DB	
QY	1581 GATGATGTGCTTGGAAAAAATCTGTATATATGATGCCATACATTTTCGAATGAAAGATAT 1640
DB	
QY	1670 TTTGGAGCTGGCCAAATTTGACATTTCAATATGGTGCAGGCTTACACCAAAAGGAGAGATCAA 1729
DB	
QY	1641 TTAGAACTAGCGAACTTGGACCTTCAATAGGTGTCAGTCTATACCAACCAACAGAGCTTCAA 1700
DB	
QY	1730 CACATTTGTCAGCTGGTGGAGAGAAATCGGATTTCAATGATCTTACATTCACCCGCCAGCGG 1789
DB	
QY	1701 GATCTTCGAAAGGTGGTGGAAATCATCCGCTTTTACCGGATCTGAAATTTCTCTGAGCGGT 1760
DB	
QY	1790 CTTGTGGAATGATTTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCTGCTTGT 1849
DB	
QY	1761 GTGACCGGAAATATATTTCTCACCGGATCCCTTTATCTTTGAGCCCGAGTTTCTAAGTGC 1820
DB	
QY	1850 AGAATTTGCTATGCCAAGACTTCTTCTGCTCGCAGTTTATTTAGACGATCTTTACGACACC 1909
DB	
QY	1821 AGAGAGTTTATACAAAATTTTCCAAATTTTCACTGTTTATTTAGATGATCTTTATGACGCC 1880
DB	
QY	1910 CACGGATCTCTGGATGATCTTAAATTTGTTCTCTGAAGCGGTCCGAGATGGGATATCTCT 1969
DB	
QY	1881 CATGGATCTTTAGACGATCTTAAATTTGTTTTCACAGAAATCAGTCAAAAGATGGGATCTATCA 1940
DB	
QY	1970 GTCTCGATAGCGTTCCGGATATCATGTTGAAAGTTTGTCTTCTAGGGCTGTACAAACACA 2029
DB	
QY	1941 CTAGTGGG---CCAAATGCCAACAAATGAATAATATGTTTGTGGGTTTCTCAATACT 1997
DB	
QY	2030 GTCAATGGGATTTGGAAAAAGATGCACTCAAGGAAACAAGCGCGTGAATGTGTGGGCTATCTT 2089
DB	
QY	1998 TTTTAATGATATACAAAAGAGGACGTCGAGAGGCAAGGGCGCGATGTGTAGGCTACATTT 2057
DB	
QY	2090 CGAAAAGTATGGAGGCGTGTCTCGCATCGTATACCAAGCAAGCCGAATGTCGCGCAGCA 2149
DB	
QY	2058 CAAAATGTTTGGAAAGTCCAACTTGAAGCTTTACCGAAAGAGCAGAAATGTCCTGAAGCT 2117
DB	

QY 2150 AAGTATGTCGCGACATTCACGAATATGTGAAATGTCGAAAGTGTCCATAGCACTTGGC 2209  
 DB |||||  
 QY 2118 AAATATGTCGCGACATTCACGAATATGTGAAATGTCGAAAGTGTCCATAGCACTTGGC 2177  
 DB |||||  
 QY 2210 ACAGTCTGCTAACTCAATCTTTTTCACCTGGAGAAATTAATCTCTGATTAATCAATTTACAG 2269  
 DB |||||  
 QY 2178 ACAGTCTGCTAACTCAATCTTTTTCACCTGGAGAAATTAATCTCTGATTAATCAATTTACAG 2237  
 DB |||||  
 QY 2270 CAAGTAGACCTTGGTCAATTTCTGATCTTGTGCTTTGATCTGAGCACTAAATCAAT 2329  
 DB |||||  
 QY 2238 AAAAATTGATCGCGAATCTAGATTTCTTCAACTCATGCGGCTTAACAGGGCGTTTGGTGAAT 2297  
 DB |||||  
 QY 2330 GACACCAAGCTTACACGAGCGGAGAGAAACCGTGTGCAATTTGGTTCCTCAGCTACAGTGC 2389  
 DB |||||  
 QY 2298 GACACCAAACTTATCAGGAGAGAGAGGTCAGGTGAGTGGCTTCTGCAATCAATGT 2357  
 DB |||||  
 QY 2390 TACATGAGGAGAAATCCGAGTGCACAGAGAGAAAGCTCTTAAGTCAATGTTTATGGPATC 2449  
 DB |||||  
 QY 2358 TATATGAAGGACCATCTAAATCTCTGAAGAGAGCTCTAACAACATGTCTATAGTGC 2417  
 DB |||||  
 QY 2450 ATCGACACCACTGAAGGAATGAATTTGGAGTGGAGTGGCCAAACCCAGGAGCAATGCCCA 2509  
 DB |||||  
 QY 2418 ATGGAATATCCCTCGAAGAGTGAATAGGAGTGTGTGAA-----TAACAAAATACCG 2471  
 DB |||||  
 QY 2510 TTGTGTGTGAGAGACTGCTGTTCAACACTGCAAGTGTGATGAGCTGTTTATATGTAC 2569  
 DB |||||  
 QY 2472 GATATTTACAAAGACTGGTTTTTGAACCTGCAAGAAATATGCAACTCTTTTATATGCAA 2531  
 DB |||||  
 QY 2570 AGAGATGGCTTGGTATCTC---TGACAAAGAGATGAAGACCATGTGACCGCAACTCTT 2626  
 DB |||||  
 QY 2532 GGGATGTTTGACACTATCAGATGATGGAATTAAGAGCATGTCAAAATTTGCTC 2613  
 DB |||||  
 QY 2627 TTCGATCTGTGGGTAGCATA 2648  
 DB |||||  
 QY 2592 TTCAACCACTGTGCTAGATTA 2613  
 DB |||||

RESULT 2

AX87530  
 ID AX87530 standard; DNA; 2525 BP.  
 XX  
 AC AX87530;  
 XX  
 DT 08-OCT-1999 (first entry)  
 XX  
 DE E-alpha-bisabolene synthase computer-generated nucleic acid sequence.  
 XX  
 KW E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;  
 KW gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour; ss.  
 OS Synthetic.  
 OS Abies grandis.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 1..2454  
 FT /\*cag= a  
 XX  
 PN WO9937139-A1.  
 XX  
 XX 29-JUL-1999.  
 XX  
 PF 21-JAN-1999; 99WO-US001300.  
 XX  
 PR 22-JAN-1998; 98US-0072204P.  
 XX  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX  
 XX Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;  
 XX WPI; 1999-458569/38.  
 XX P-PSDB; AAY06567.  
 XX  
 XX Grand fir (Abies grandis) sesquiterpene synthase genes and protein

PT products, useful for sesquiterpenoid production.

XX Example 11; Page 95-98; 136pp; English.

XX This is the sequence of a computer-generated nucleic acid coding for a grand fir (Abies grandis) E-alpha-bisabolene synthase polypeptide (see AAY06567) that has conservative amino acid substitutions relative to the native E-alpha-bisabolene synthase sequence given in AAY06562. E-alpha-bisabolene synthase is a wound-inducible enzyme capable of generating multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl diphosphate. The invention provides gymnosperm sesquiterpene synthase nucleic acids and recombinant proteins, as well as expression vectors, and a method of imparting or enhancing production of a gymnosperm sesquiterpene synthase in a prokaryotic or eukaryotic host cell, especially in a plant, in order to impart, enhance or modify production of sesquiterpenoids, and hence to alter the flavour or fragrance of plant products, to improve disease resistance or to alter ecological interactions mediated by farnesyl diphosphate and its derivatives

SQ Sequence 2525 BP; 749 A; 498 C; 600 G; 678 T; 0 U; 0 Other;

Query Match 27.0%; Score 731.6; DB 2; Length 2525;  
 Best Local Similarity 59.7%; Pred. No. 3.6e-187;  
 Matches 1369; Conservative 0; Mismatches 894; Indels 31; Gaps 7;

QY 434 AATCCATCTGCATATGATACAGCTTGGTGGCAAGATTCGGTCAATTTAGCGCTCTGGT 493  
 DB |||||  
 QY 238 ACTCCATCTCTTATGACACAGATGGGTAGCAGGGTGCCGCCATTTGATGGCTCTGCT 297  
 DB |||||  
 QY 494 GCACCCCAATTTCCCAACCGCTTCAATGGATTTCTGAACAATCAACTGCCAGATGGCTCG 553  
 DB |||||  
 QY 298 CGCCCGCAATTTCCCAACCGATTTGATGGATTTGAAAACCAAGTTAAAAGATGGTTCA 357  
 DB |||||  
 QY 554 TGGGGTGAGGAGTGCAATTTTCTGGCGTATGACAGAGATTTTAAACACTCTCGCTGCTC 613  
 DB |||||  
 QY 358 TGGGGAAATTCAGTCCCACTTTCTGCTGTCGACCGTCTTCTTGCACTCTTCTTGTGTT 417  
 DB |||||  
 QY 614 CTCACCTCTCAAAATATGGAATAGGGGCAATTTCAAGTCAGAGAGGGTTGATTTGTG 673  
 DB |||||  
 QY 418 CTTGTGCTCTTAAATGGAAACGTTGGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA 477  
 DB |||||  
 QY 674 AGAAAACACATGGAAGAAATGAAGGACGAGCTGCAATCACAAG---GCCAAGTGAATTC 730  
 DB |||||  
 QY 478 AAGAGCAATCTGGAACCTAGTAAAGGATGAACCGATCAAGTAGCTTGGTAAACAGACTTT 537  
 DB |||||  
 QY 731 GAGTCTGTTTCTCTCAATTTAGATGAAGCAAAAAGCTTGGGATTTGGATCTTCTTAT 790  
 DB |||||  
 QY 538 GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC 597  
 DB |||||  
 QY 791 CACTCTCTTCTCAATCTCCCAATCCCAACCAAGCGGCGAGAAAAGCTTCAAAAGATTTCCC 850  
 DB |||||  
 QY 598 GACCTGCTTTATATACATCTGTTGCAAGCTTAAACGGCAGGAAAGATTAGCANAACCTTCA 657  
 DB |||||  
 QY 851 CTCAAATGTTCTTCAATACCAATCAGAGCGGTCTCTACTCTCTGAGAGGTTTGCAGAGT 910  
 DB |||||  
 QY 658 AGGAGGAATTTATGCGGTTCCGTGCCATTTGTTATTTCTTTAGAGGAATACAAGAT 717  
 DB |||||  
 QY 911 GTGGTGGACTGGCAAGAGATCAAAAATCTTCAATCAAGAGACGGATCAATTTTAAAGTCC 970  
 DB |||||  
 QY 718 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAGTCAGGATGGTCTTTCTTAAAGCTCA 777  
 DB |||||  
 QY 971 CTTGCAATCTACTGCTGTTGTTTCAATGCACTCAAAAACCAACGATGCTTCCCTCTCTC 1030  
 DB |||||  
 QY 778 CTTGCTTCTACTGCTGCTGTTTTCATGCAACAGGAGACCGGAAATGCTTGAATTTCTTG 837  
 DB |||||  
 QY 1031 AACTCTGCTCAGCAAAATTTGGCGACTAGTCTTCTTCCATTTACCCACTTGATCTATTT 1090  
 DB |||||  
 QY 838 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGCTATCTCTGAGATCTGCTG 897  
 DB |||||  
 QY 1091 GAACGCTCTGGGCTGTGATACAGATTTGAACCGCTTGGGAATCGATCTGATTTTCAAGAAA 1150  
 DB |||||  
 QY 898 GAACGCTGTGATCGTAGATATATTTGACGCTTGGAAATCTATAGACACTTTGAAAG 957  
 DB |||||



1151 GAAATCAAAGAATCTCTGATTTACGTTTATAGGTACTGGAGCCGAAAGGCGTGGGA 1210  
1152 |||||  
958 GAAATCAAGGAAGCTCTTGAATATGTTTACAGGCATTTGGAACG---AAAGAGGAATGGG 1014  
1159 |||||  
1211 TGGCAAGATGCAATCTCTATTCCTGATGTCGATGACACTGCCATGGGTCTTGAATCCTG 1270  
1212 |||||  
1015 TGGGCGAGCTAAATCCCNACAGCATCTTGAGACCACTGCTTTGGGATTTGATGCTT 1074  
1216 |||||  
1271 AGACTTCATGGAATCAATGATGTTCTTCAAGATGTTCTGGAGAAATTTGAGAGCAGAGAAAGGA 1330  
1272 |||||  
1075 CGGCTGCATAGTACAATGATCTCCAGCCATTTTGGACCACTTTGACCACTTCAAAAGATGCCAATGGG 1134  
1276 |||||  
1331 GACTTCTTTTGGCTTGGCGGTCAACGCAAAATGTTGTTGACCGATAATCTTAACCTTTAT 1390  
1332 |||||  
1135 AAATTCATTTGCTGCACCGGTCAATTCACAAAGATGTTAGCAAGCATGCTGAACTTTAT 1194  
1336 |||||  
1391 AGATGTTCAACAGTATGTTTCCGGGAGAAAAAGATAATGGAAGAACTTAAGACCTTCACT 1450  
1392 |||||  
1195 AGAGCTTCCCAGCTCGCATTTCCCGGAGAAACATTTCTTGATGAAGCTAAAGCTTCGCT 1254  
1396 |||||  
1451 ACAATCATCTCAAAATGCTCTTGGCAAAAACAAGCAATTTGATAAGTGGGCTGTCAAG 1510  
1452 |||||  
1255 ACTAAATATTTGAGAGAACTCTTGAGAAAGTGAAGCTTCCAGTGCATGGAACAACAA 1314  
1453 |||||  
1511 AAGGATCTCTCGGAGAGGTGGAGTATGCTATAAGTATCCGTGGCATAGAAATGATGCCA 1570  
1512 |||||  
1315 CAAAACCTGAGCAGAGATCAATACGCGCTGGAAGCTTCTTGGCATGCCAGTGTCCG 1374  
1513 |||||  
1571 AGATTGGAGGCAAGAGTATACATAGAGCAATTTGGATCAAAATGATGTTGGCTGGGGAAG 1630  
1572 |||||  
1375 AGAGTGAAGCAAGAGATGCTCAAGTGTATCGCCAGATTTATGACGCTATGCAACAA 1434  
1573 |||||  
1631 ACTGTGATAGATGCTATATGTGACACGCAAAATATTTGGAGCTGGCCAAATTTGGAC 1690  
1632 |||||  
1435 TCGGTTTACAGCTACCTTACCTTTCGCTCGGAGAGGCGCTGGAATTTACTTTCTTA 1494  
1633 |||||  
1691 TTCATATGCTGAGCGCTTTACACCAAAAGAGACTCAACACATTTGCTAGCTGGTGAGA 1750  
1692 |||||  
1495 TTCACATTTCCNGTCCATCCACCAAGAAAGAAATGAAGATGTTACCAGCTGGTTTGA 1554  
1693 |||||  
1751 GAATCGGGAATCAATGATCTTACATTCACCGCCAGCGGCTGTGGAATGATTTCTCA 1810  
1752 |||||  
1555 GATTCGGGGTGGCACTTATTCACCTTTCGCTCGGAGAGGCGCTGGAATTTACTTTCTTA 1614  
1753 |||||  
1811 GTGGCGGTAGTATGTTGAGCCAGATTCGCTGCTGTGATGATTCCTATGCAAGACT 1870  
1812 |||||  
1615 GTAGCGCGGGACCTATGAAACCCAGTATGCCAAATGCAAGTTCCTCTTTACAAAGTG 1674  
1813 |||||  
1871 TCTTGCTCGCAGTTATTTCTAGACGATCTTTACGACACCCACCGATCTCTGGATGATCTT 1930  
1872 |||||  
1675 GCATGCTTCGAGACTGTTCTGGAGCATATGATGACACTTATGGAACCCCTAGATGAATG 1734  
1873 |||||  
1931 AAATGTTCTGGAAGGTCGGAAGTGGATATCTCTGTGCTGATAGCTGCGGTTCGGAT 1990  
1932 |||||  
1735 AAGCTATTCATGAGGCTGTGAGAAGATGGGACCTCTCTCTTTACAGAAACCTTCCAGAC 1794  
1933 |||||  
1991 AATCAGTTGAAGTGTGCTTCTTAGGCTGTACAAACACAGTGAATGGAATTTGAAAAGAT 2050  
1992 |||||  
1795 TAT---ATGAATCTATGTTACCAATCTATATGACATGTTTACGAGGTGGCTGGAG 1851  
1996 |||||  
2051 GGAATCAAGGAACAGCGCTGATGCTGGGCTATCTTCGAAAAGTATGGGAGGGGCTTG 2110  
2052 |||||  
1852 GCAGAGAAGGAACAGGGGCGTGAATTTGGTCAAGCTTTTTCAGAAAGGATGGGAGGATAT 1911  
2053 |||||  
2111 CTGCTATGATACCAAGAGCGGATGCTGGCGAGCAAGTATGTCGCACTTCAAC 2170  
2112 |||||  
1912 CTTCTGGGTTTATATGAAGAGCTGAATGTTAGTCTGCTGATGATGTCCTTCTTGGAC 1971  
2113 |||||  
2171 GAATATGTGGAATGCAAAAGTGTCCATAGCACTTTCGCAAGTGTGTAATACTCAATC 2230  
2172 |||||  
1972 GAGTACATAGAATGAATGGAATCACAATCTATCGGCCCAAGTATCTTCTGTTGAGTGGAGTG 2031  
2173 |||||  
2231 TTTTTC---CTGAGGAATTACTCTCTGATATACATTTTACAGCAAGTAGACCTTCGGTCC 2287

2032 TTGATAATGGATGGCACTCTTTTCGCAAGAGGCATTTAGAGAAAGTAGATTTATCCAGGA 2091  
2033 |||||  
2288 AAATTTCTGC-----ATCTTGTGCTTTGACTGCAAGCACTAATCAATGACACCAAGACT 2341  
2034 |||||  
2092 AGACGTTTCTCACAGAGCTGAATAGCCTCATTTCCCGCCTGGCGGATGACACGAGACA 2151  
2035 |||||  
2342 TACCAGGCGGAGAGAAACCGTGGTGAATTTGGTTCAGCGTACAGTGTCTACATGAGGGA 2401  
2036 |||||  
2152 TATAAGCTGAGAGGCTCGTGGAGAAATTTGGGTCAGCATTTGAATGTTTACATGAAAGAC 2211  
2037 |||||  
2402 AATCCGAGTGCACAGAGAAAGAACTCTAAGTCTATGTTTATGTTATGTTATCATCATCAACGCA 2461  
2038 |||||  
2212 CATCTGAAATGACAGAGAAAGGCTCTCGATCATCATCTATAGCATTTCTGAGCGCGG 2271  
2039 |||||  
2462 CTGAGAGGAAATGAATTTGGAGTTGGCCCAACCCAGCGCAATGCCCATTTGTGTGAGA 2521  
2040 |||||  
2272 GTGAAGGAACTGACAAAGAGAGTTTCTGAAGCCCG---AGAGCTCCCATTTCCGCTGCAAG 2328  
2041 |||||  
2522 AGACTGCTGTTTCAACACATGCAAGAGTGTGATGCAAGCTGTTTATATGATGACAGATGCTTT 2581  
2042 |||||  
2329 AGATGCTTTTCGAGGAGACAAAGTGTGAGTGTGATATTCAGGATGGAGATGGATTC 2388  
2043 |||||  
2582 GGTATCTCTGACAAAGAGATGAAGACCATGTGACGCCGAACTCTTTTCGATCTGTGCGG 2641  
2044 |||||  
2389 CGTGTGTTCCAAATTAGAAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCCA 2448  
2045 |||||  
2642 TAGCA-----TACTGATATATATATATATTTCAATTTCAATCCAAACCAAAAAA 2691  
2046 |||||  
2449 CTGTAATCAAAATGTTGCAATATATTTGAATATCAACTATGTTTTCACAAAAA 2508  
2047 |||||  
2692 AAAAAA 2705  
2048 |||||  
2509 AAAAAA 2522

RESULT 3  
AA08654  
ID AA08654 standard; cDNA; 2424 BP.  
XX  
AC AA08654;  
XX  
XX 27-SEP-1999 (first entry)  
XX  
DE Grand Fir monoterpene synthase clone AG1.28.  
XX  
XX Myrcene synthase; limonene synthase; pinene synthase; flavour;  
KW monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;  
KW probe; ss.  
XX  
OS Abies grandis.  
XX  
XX Key Location/Qualifiers  
FT 2..2350  
FT /\*tag= a  
FT /product= "Potential diterpene cyclase"  
XX  
XX WO9902030-A1.  
XX  
XX 21-JAN-1999.  
XX  
XX 10-JUL-1998; 98WO-US014528.  
XX  
XX 11-JUL-1997; 97US-0052249P.  
XX  
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
XX Bohlmann J, Steele CL, Croteau RB;  
XX  
XX WPI; 1999-120396/10.  
XX  
XX P-PSDB; AAW85703.  
XX  
XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand

fir (*Abies grandis*), used to provide plants with modified production of monoterpenes, e.g. myrcene, limonene or pinene.

Example 1; Page 86-90; 121pp; English.

Nucleotide sequences encoding myrcene synthase, limonene synthase and pinene synthase from Grand fir may be incorporated into any organism (e.g. intact plant, animal, microbe), or derived cell culture that produces geranyl diphosphate for the production of the aforementioned enzymes or their products. The sequences when expressed in transfected cells may also be used for the production or modification of flavour and aroma properties, improvement of defense capability, and the alteration of other ecological interactions mediated by myrcene, limonene, pinene, or their derivatives. In particular they can be used for the production of plant seeds for the extraction of oil or meal. cDNAs encoding myrcene synthase, pinene synthase and limonene synthase from Grand Fir were isolated and sequenced by the following procedure. Based on comparison of sequences of different plant synthase enzymes, four conserved sequences were identified for construction of four degenerate primers (AAx08646-49). One of these primers (AAx08648) was an antisense primer and used for PCR with the other three sense primers. Only a combination of primers AAx08648 and AAx08649 generated a PCR product which was 110 base pairs in length. The amplified product was used in a plasmid construct to transform *E. coli* XL1-Blue cells and the inserts were sequenced. Four probes (AAx08650-53) were devised from the sequenced insert and were used to screen a Grand Fir cDNA library. One of these clones is the myrcene synthase sequence given in AAx08643. The other clones identified are described in AAx08654-56

Sequence 2424 BP; 730 A; 472 C; 575 G; 647 T; 0 U; 0 Other;

Query Match 27.0%; Score 730.2; DB 2; Length 2424;

Best Local Similarity 59.6%; Pred. No. 8.5e-187;

Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;

```
Qy 434 AATCATCTGATATGATACAGCTTGGGTGGCAAGAAATTCGTCAAATGACGGCTCTGGT 493
Db 134 ACTCCATCTGTTATGACACAGATGGGTAGCGAGGGTGCCCGCATGTGATGGCTCTGCT 193
Qy 494 GCACCCCAATTTCCCAAAACGCTTCAATGGATTTCTGAACAACTCAACTGCCAGATGGCTCG 553
Db 194 CGCCCGCAATTTCCCAAAACAGTTGACTGGATTTTGAACCACTGTTAAAGATGGTTCA 253
Qy 554 TGGGGTGAGGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACATCTCGCTGCTCCTC 613
Db 254 TGGGGAAATTCAGTCCCACTTCTCTGTCTCCGACGCTCTTCTTGCCACTCTTCTTGTTGT 313
Qy 614 CTCACTCTCAAAATATGGNATAAGGCGCAGCATTTCAAGTCGAGAGAGGGGTTGATTTGTG 673
Db 314 CTTGTGCTCTTAAATGGAAAGCTTTGGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA 373
Qy 674 AGAAAAACACATGGGAAGAAATGAAGGACGAGCTGACAAATCACAG--GCCAAGTGGATTC 730
Db 374 AAGAGCAATCTGGAACATAGTAAGAGATGAAACCGATCAAGTAGCTTGGTAACAGACTTT 433
Qy 731 GAGTCGTGTTTCTGCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGATCTTCTTTAT 790
Db 434 GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCCTCGGACTTCCCTAC 493
Qy 791 CACCTCCCTTTCACTCTCCCAATCCACCAAGCGCCAGAAAAGCTTCAAGAGTTCCC 850
Db 494 GACCTGCTTATATACATCTGTTGACACTTAACCGGAGGAAGATTTAGCAAAACTTTCA 553
Qy 851 CTCAATGTTCTTCAATACCATCAGACGCGTTGCTCTACTCTCTGAGGGGTTTGCAGAAT 910
Db 554 AGGAGGAAATTTATGCGGTTCCGTCGCCATTTGTTATTTCTTTAGAGGAAATACAGAT 613
Qy 911 GTGGTGACTGGCAAGAGATCAAAATCTTCAATCAAGACCGGATCATTTTAAAGTCC 970
Db 614 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAAGTCAGGATGGGTCTTCTTAAAGCTCA 673
Qy 971 CTTGCATCTACTGTTGTTGTTCTTATGCACACTCAAAACAAACGATCGCTCCACTTCTC 1030
Db 1030
```

```
Db 674 CTTGCTCTACTGCTCGCTGTTTTTCATGCACACAGGAGACGCGAAATGCTTGAATCTTG 733
Qy 1031 AACTTCGTGCTCAGCAAAATTTGGCGACTAGCTTCTTTGCCATTTACCCACTTGAFTATTT 1090
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGCTATCTCTGTGGATCTGCTG 793
Qy 1091 GAACGGCTCTGGGCTGCGATACAGTTGAACGCTTGGGATCGATCGCTATTTCAAGAAA 1150
Db 794 GAACGGCTGTTGATCGTAGATAAATTTGTACGGCTTGGAAATCTATAGACACTTTTGAAG 853
Qy 1151 GAAATCAAGAAATCTCTCGAATTACGTTTATAGGTACTTGGGACGCGCAAGAGCGCTGGGA 1210
Db 854 GAAATCAAGAAAGCTCTTGATTTATGTTTACAGGCATTTGGAACG--AAAGGAATTTGGG 910
Qy 1211 TGGGCAAGATGCAATCTATTTCTGATGTCATGACACTGCCATGGGTCTTTAGAAATCTG 1270
Db 911 TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTTGGGATTTTCGATTCCT 970
Qy 1271 AGACTTCATGATACAAATGATCTTTCAGATGTTTCTGGAGAAATTTTCAGAGACGAGAAAGGA 1330
Db 971 CGGCTGCATAGGTACAAATGATCTTCAGGCCATTTTTCAGCAACTTTCAAGAGATGCCAATGGG 1030
Qy 1331 GACTTCTTTTGTCTTGGCGGTCAAAACGCAAAATTTGGTGTGACCGATATCTTTAACTTTAT 1390
Db 1031 AAATTCATTTGCTCGACGGGTCAATTCACAAAGATGTAGCAAGCATGCTGATCTTTAT 1090
Qy 1391 AGATGTTTCAACATATGTTTTCGGGAGAAAGATTAATGGAAGAGCTTAAGACCTTCACT 1450
Db 1091 AGAGCTTCCCGAGCTCGCATTTTCCGGAGAAACATTTCTTGATGAAGCTTAAAGCTTCGCT 1150
Qy 1451 ACAATATCTTCCAAATGCTTCTTGCCTCAAAACCAACGCAATTTGATAGTGGGCTGTCAAG 1510
Db 1151 ACTAAATATTTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAACCAACAAA 1210
Qy 1511 AAGGATCTTCTCGAGAGGTGGAGTATGCTATATAAGATATCCGTGGCATAGAGATATGCCA 1570
Db 1211 CAAAACCTGAGCCCAAGAGATCAATATACGCGCTGAGACTTCTTGGCATGCGAGTGTCCG 1270
Qy 1571 AGATGGAGGCAAGAGTATACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG 1630
Db 1271 AGAGTGGAGCAAAAGAGATCTGTCAAGTGTATCGCCAGATTTATGACGCATAGCAAAA 1330
Qy 1631 ACTGTGTATAGATGCTATATGTGAGCAAGCAAAATATTTGGAGCTGGCCAAATTTGGAC 1690
Db 1331 TGGGTTTACAAAGCTACCTTACGTTGAAACAAATGAAAGGTTTTTATAGAGCTGGGAAAATTAGAT 1390
Qy 1691 TTCAAATATGTGCGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTGCTGGTGAGGA 1750
Db 1391 TTCAAATATCCAGTCCATCCACCAAGAGAAATGAAAGATGTTACCAGCTGGTTTAGA 1450
Qy 1751 GAATCGGGATTTCAATGATCTTACATTTCACTCCCGCAGCGGCTGTGGAAATGTATTTCTCA 1810
Db 1451 GATTCGGGTTGCCACTTATTCACCTTCGCTCGGAGAGGCGCTGGAATTTCTACTTCTTA 1510
Qy 1811 GTGGCGTTAGTATGTTTGGCCAGAAATTCGCTGCTTGTAGAAATTTGCTATGCCAAGACT 1870
Db 1511 GTAGCGCGGGGACCTATGAACCCAGATATGCCAAATGCGAGGTTCTCTTTTACAAAAGTG 1570
Qy 1871 TCTTGCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCAGCATCTCTGGATGATCTT 1930
Db 1571 GCATGCTGACAGACTGTTCTGGAGCATATGTATGACACTTATGGAACCTTAGATGAATTTG 1630
Qy 1931 AAATTTGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGTGGATAGCGTTTCGGGAT 1990
Db 1631 AAGCTATTCACTGAGGCTGTGAGAAGATGGGACCTCTCTCTTTTACAGAAAACCTTCCAGAC 1690
Qy 1991 AATCAGTTGAAAGTTTCTTCTAGGGCTGTACAAACAGTGAATGGAATTTGGAAGAT 2050
Db 1691 TAT---ATGAAACTATGTTTACCAAAATCTATTATGACATAGTTTACGAGGTGGCTTGGAG 1747
Qy 2051 GGACTCAAGCAACAAAGCGCTGATGCTGGCTATCTTCGAAAAGTATGCGAGGGCTTG 2110
Db 1748 GCAGAGAAAGAACAGGGGCGTGAATTCGTGCTGCTTTTTCAGAAAGGATGGAGGATAT 1807
```



```
QY 851 CTCAATGTTCTTCATACCACATCAGACGGCGTGTCTACTCTCTGAGGGTTCAGAGAT 910
Db 554 AGGAGAGAAATTTATGCGGTTCCGTCGCCATTTGTGTAATCTTTAGAGGGAATACAGAT 613
QY 911 GTGTGACTGGCAAGAGATCACAATCTTCAATCAAGAGACGAGATCAATTTTAAAGTCC 970
Db 614 ATAGTTGAAATGGGNAACGAATAATGGAAGTTCAAGTCAGAGTGGTCTTCTTAAGCTCA 673
QY 971 CTGCACTACTGCTTGTGCTTCATGCACTCAAAACAAACGATGCCCTCCACTTTCTC 1030
Db 674 CTTGCTTTCTACTGCTCGCTTTTCATGCACACAGGAGACGGAATGCCCTTGAATCTTG 733
QY 1031 AACTTGTGTCTCAGCAATTTGGCGACTAGCTTCTTCCCTTGGCAATACCACTTGTACTATTT 1090
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTTGGCTGTATCTCTGTGGATCTGCTG 793
QY 1091 GAACGCTCTGGGCTGTGCGATCAGATTGAACGCTTGGGAATCGATCGCTATTTCAAGAAA 1150
Db 794 GAACGCTGTGTAGTAGATATAATTTGACGCTTGGAAATCTATAGACACTTTGAAAG 853
QY 1151 GAAATCAAGAAATCTCTGGATTAAGTTTATAGTFACTGGGACGCGGAAAGAGCGGTGGGA 1210
Db 854 GAAATCAAGGAAGCTCTTGATATATTTTACAGGCAATTTGGAACG---AAGAGGAATTTGGG 910
QY 1211 TGGGCAAGATGCAATCTTATCTGATGTGATGACACTGCCATGGTCTTAGAATCCTG 1270
Db 911 TGGGCGAGACTAAATCCATAGCAGATCTTTGAGACCACTGCTTTGGGATTTGATGCTT 970
QY 1271 AGACTTCATGATCAATGATCTTTCAGATGTTCTGAGAAATTTCAAGACGAGAAAGGA 1330
Db 971 CGGCTGCTAGGTACATGATGATCTCCAGCCATTTTGGACAACTTCAAGATGCCATGGG 1030
QY 1331 GACTTCTTTTCTGTTGCGGTCAAAACGCAAAATTTGGTGTGCCGATTAATCTTAACTTTAT 1390
Db 1031 AAATTCATTTGCTCGACCGGTCAATTCACAAGAGATGTACAAAGCATGCTGAATCTTTAT 1090
QY 1391 AGATGTTCAAGTATGTTTTCGGGAGAAAGATTAATGGAAGAGCTAAGACTTCACT 1450
Db 1091 AGAGCTTCCCAGCTCGCAATTTCCGGGAGAAACAACTTCTTGATGAAGCTAAAGCTTCGCT 1150
QY 1451 ACAATCATCTCCAAATGCTCTTGCCAAACAAACGCAATTTGATAGTGGGCTGTCAAG 1510
Db 1151 ACTAATAATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGATGGAGAACACAA 1210
QY 1511 AAGGATCTTCTGGAGAGTGGAGTATGCTATAAGTATCCGTGGCAGTAGAAGTATGCCA 1570
Db 1211 CAAAACTGAGCCAAAGAGATCAAAATACGCTGAAGACTTCTTGGCATGCCAGTGTCCG 1270
QY 1571 AGATTTGAGGCAAGAGTATACATAGACAAATTTGGATCAAAATGATGTCTGCTGGGGAAG 1630
Db 1271 AGAGTGGAAAGCAAGAGATACTGTCAAGTGTATCGCCCAAGATTTATGCACGATAGCAAAA 1330
QY 1631 ACTGTGTATAAGTCTATATGTAGCAACGAAATATTTGGAGCTGGCCAAATTTGGAC 1690
Db 1331 TGGCTTTTCAAGCTACCTACGTGAAACATGAAAGTGTTTTAGAGCTGGGAAATTAGAT 1390
QY 1691 TTCAATATGTTGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCACTGGTGAGA 1750
Db 1391 TTCAACATTTCCAGTCCATCCACCAAGAGAAATGAAGATGTTACCACTGGTTTAGA 1450
QY 1751 GAATCGGATTTCAATGATCTTACATTCACCGCCAGCGGCTGTGGAAATGTATTTCTCA 1810
Db 1451 GATTCGGGGTTGCCACTATTTTCACTTCGCTCGGAGAGGCGCTGGAAATTTCTACTTCTTA 1510
QY 1811 GTGCGGTTAGTATGTTTGGACGAGATTCGCTGTGTAGAAATTTGCTATGTCAGACT 1870
Db 1511 GTAGCGCGGGGACCTATGAACCCAGTATGCCAAATGCAAGGTTCTCTTTTACAAAAGTG 1570
QY 1871 TCTTGCTTCGAGTTATTTCTAGACGATCTTTTACGACACCCACCGATCTCTGGATGATCTT 1930
Db 1571 GCATGCTTGACAGCTGTTCTGACGATATGATGACACTTATGGAACCTTAGATGATG 1630
QY 1931 AAATTTGTTCTGGAACGGTCCGAAGATGGGATATCTCTGTGTGGATAGCGTTCGGGAT 1990
```

```
Db 1631 AAGCTATTTCACTAGGCTGTGAGAAGATGGGACCTCTCTTTTACAGAAAACCTTCCAGAC 1690
QY 1991 AATCAGTTGAAAGTGTTCCTTAGGGCTGTACAAACACAGTGAATGATTTGGAAAAAGAT 2050
Db 1691 TAT---ATGAAACTATGTTTACCAAACTATTATGACATAGTTCACGAGGTGGCTTGGAG 1747
QY 2051 GGACTCAAGGAACAAGCCGCTGATGTGCTGGGCTATCTTCGAAAAGTATGAGAGGGCTTG 2110
Db 1748 GCAGAGAAGGAACAGGGGGCTGAAATTTGGTCAGCTTTTTCAGAAAGGATGGAGGATTAT 1807
QY 2111 CTCGCATCTGATATACCAAGAACCGGAATGCTGCGCAGCAAGTATGTGCCGACATTCAC 2170
Db 1808 CTTCTGGTTTATGAGAAGCTGAAATGTTAGTCTGCTGAGTATGTGCTTACCTTGAC 1867
QY 2171 GAATATGTGGAAATCCAAAGTGTCCATAGCACTTTGCGACAGTCTGTAATACTAACTCAATC 2230
Db 1868 GAGTACATAAAGAAATGGAATCACTATCTGCGCAACGTATACTTCTGTTGAGTGGAGTG 1927
QY 2231 TTTTTC---CTGAGAAATTACTTCTCTGATTAATTTTACAGCAAGTAGACTTTCGGTCC 2287
Db 1928 TTGATTAATGGATGGGCAACTCTCTTTCGCAAGAGGCATTTAGAGAAAGTAGATTATCCAGGA 1987
QY 2288 AATTTTCTGC-----ATCTTGCTCTTTGACTGAGCACTAATCAATGACACCAAGACT 2341
Db 1988 AGACGTGTTCTCAGAGCTGAAATAGCCTCATTTCCCGCTGGCGGATGACACGAAGACA 2047
QY 2342 TACAGCGCCAGAGAAACCGTGTGAAATTTGCTTTTCCAGCGCTACAGTGTCTACATGAGGGAA 2401
Db 2048 TATAAAGCTGAGAAGGCTCGTGAGAAATTTGGCGTCCAGCAATTGAAATGTTACATGAAGAC 2107
QY 2402 AATCCGAGTGACACAGAGGAAGCTCTAAGTCAATGTTTATGTTATCATCGATCATCGACAA 2461
Db 2108 CATCTGAATGTATACAGAGGAAGAGGCTCTCGATCACAATCTATAGCAATTTCTGGAGCCGCG 2167
QY 2462 CTCAAGGAATGAATTTGGGAGTTGGCCACCCAGCGAGCAATGCCCATTTGTGTGTGAGA 2521
Db 2168 GTGAAGGAATGACAAGAGAGTTCGAAAGCCG---ACGACGTCCCATTTGCGCTGCAAG 2224
QY 2522 AGACTGCTGTTTCAACACTGCAAGAGTGTATGACGTGTTTATATGTACAGAGATGGCTTT 2581
Db 2225 AAGATGCTTTTCGAGGAGACAAGTGTGATATTTCAAGGATGAGATGGATTC 2284
QY 2582 GGTATCTCTGACAAAGAGATGAAAGACCATGTGACCGGAACTCTTTTTCGATCTCTGTGGCG 2641
Db 2285 GGTGTTTCCAAATTAGAAGTCAAGATCATATCAAAGAGTGTCTCATTTGAACCGCTGCCA 2344
QY 2642 TAGCA-----TACTGATATATATATATATATATATATCAATTCATCAATCCAAAAAAA 2688
Db 2345 CTGTAATCAAAATGTTGCAATTAATAATTGAAATAATGTCAACTATGTTTTCACAAAAAA 2404
QY 2689 AAAAAAAAAAAAAAAAAA 2705
Db 2405 AAAAAAAAAAAAAAAAAA 2421
```

## RESULT 5

AAAF73382

ID AAF73382 standard; cDNA; 2424 BP.

XX AAF73382;

XX XX

DT 30-APR-2001 (first entry)

XX

DE Grand fir abietadiene synthase coding sequence fragment SEQ ID NO: 15.

XX

KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;

KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;

KW terpinolene synthase; insect resistance; nutrition; ss.

XX

OS Abies grandis.

XX

PN W0200107565-A2.

XX	01-FEB-2001.				
XX	PD				
XX	PF	24-JUL-2000; 2000WO-US020264.			
XX	PF	26-JUL-1999; 99US-00360545.			
XX	PA	(UNIW ) UNIV WASHINGTON STATE RES FOUND.			
XX	PI	Steele CL, Bohlmann J, Croteau RB, Phillips MA;			
XX	PI	WPI; 2001-182782/18.			
XX	DR	P-PSDB; AAB69373.			
XX	PT	New nucleic acid encoding monoterpene synthases, for increasing terpene			
XX	PT	synthesis in plants, e.g. for increasing resistance to pests or for			
XX	PT	treatment of cancer.			
XX	PS	Example 1; Page 119-122; 175pp; English.			
XX	CC	The present invention provides the protein and coding sequences of			
XX	CC	monoterpene synthases from the grand fir. These include (-)-camphene			
XX	CC	synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-			
XX	CC	limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase			
XX	CC	and pinene synthase. The sequences can be used to produce transgenic			
XX	CC	plants expressing high levels of the enzymes, resulting in levels which			
XX	CC	are useful in protecting against and treating cancers, and to confer			
XX	CC	insect resistance on plants			
XX	SQ	Sequence 2424 BP; 730 A; 472 C; 575 G; 647 T; 0 U; 0 Other;			
	Query Match	27.0%; Score 730.2; DB 4; Length 2424;			
	Best Local Similarity	59.6%; Pred. No. 8.5e-187;			
	Matches 1370; Conservative	0; Mismatches 893; Indels 34; Gaps 7;			
Qy	434	AATCCATCTGCATATGATACAGCTTGGTGGCAAGAAATTCGCTCAATTTGACGGCTCTGGT	493		
Db	134	ACTCCATCTGCTTATGACACAGCATGGGTAGCGAGGGTCCCGCCATTTGATGGCTCTGCT	193		
Qy	494	GCACCCCAATTTCCCAACCGCTTCAATGGATTTCTGAACAACTCAACTGCCAGATGGCTCG	553		
Db	194	CGCCCGCAATTTCCCAACAGCTTGACTGGATTTTGA AAAAACCAAGTTAAAAGATGGTTCA	253		
Qy	554	TGGGGTGAGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCTC	613		
Db	254	TGGGGTAATTCAGTCCCACTTTCTGCTGTCGACCGCTCTCTTGGCCACTCTTCTTGTTGT	313		
Qy	614	CTCACTCTCAAAATATGGAATAGGGCGACATTTCAAGTGCAGAAAGGGTTGAGTTTGTG	673		
Db	314	CTTGTGCTCTTAAATGGAACGTTGGGGATCTGCAAGTAGAGCGGGAATTTGAATTCATA	373		
Qy	674	AGAAAACACATGGAAGAAATGAAGGACGAGCTGCAATCAAG---GCCAAGTGGATTC	730		
Db	374	AAGAGCAATCTGGAATAGTAAAGGATGAACCGATCAAGATAGCTTGTGTAACAGACTTT	433		
Qy	731	GAGTGTGTTTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATTTCTCTTAT	790		
Db	434	GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTTAC	493		
Qy	791	CACCTCCCTTTCTATCTCCAAATTCACCAAAAGCGCCAGAAAAGCTTCAAAAGATTTCCC	850		
Db	494	GACCTGCTTATATACATCTGTTGCAAGTAAACGGCAGGAAGATTTAGCAAACTTTCA	553		
Qy	851	CTCAATGTTCTTCATTAACCATCAGACGGGTGCTCTACTCTCTCGAGGGTTTGAAGAT	910		
Db	554	AGGAGGAAATTTATGCGGTTCCGTCCCAATTTGTTATTTCTTTAGAGGGAATACAGAT	613		
Qy	911	GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGACGGGATCATTTTTAAGCTCC	970		
Db	614	ATAGTTGAATGGGAACGAATATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTTAAGCTCA	673		
Qy	971	CCTGCATCTACTGCTGTGCTTTCATGCACTCACTCAAAAACAAACGATGCTCTCACTTCTC	1030		

QY 2111 CTCGATCGTATACCAAGAGCCGAATGGTGGCAGCAAAAGTATGTGCCGACATTCAC 2170  
 Db |||||  
 QY 1808 CTTTGGGTATATGAAGAAGCTGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1867  
 Db |||||  
 QY 2171 GAATATGTGGAAATGCCAAAGTGTCCATAGCACTTGGCAGCTGGTACTAACTCAATC 2230  
 Db |||||  
 QY 1868 GAGTACATAAGATGGATACATCTATCGCCACGATATCTCTCTGTGGTGGAGTG 1927  
 Db |||||  
 QY 2231 TTTTTC---CTGGAGAAATCTCTCTGATTAACATTTTACAGCAAGTAGACCTTGGTCC 2287  
 Db |||||  
 QY 1928 TTGATAATGGATGGCAACTCTCTTTCGCAAGAGGCAATTAGAGAAAGTAGATTATCCAGGA 1987  
 Db |||||  
 QY 2288 AAATTTCTGC-----ATCTTGTCTTTGACTGGACGCACTAATCAATGACACCAAGACT 2341  
 Db |||||  
 QY 1988 AGACGTGTCTCACAGAGCTGAATAGCCCTCATTTCCCGCCTGGCGGATGACACGAAGACA 2047  
 Db |||||  
 QY 2342 TACCAAGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGTACATGAGGAA 2401  
 Db |||||  
 QY 2048 TATAAGCTGAGAGGCTGTGGAGAAATGGCTCCAGCATTTGAATGATTAATGAAGAC 2107  
 Db |||||  
 QY 2402 AATCCGAGTGCACAGAGAAAGCTCTAAGTCAATGTTTATGTTATCATCGACACGCA 2461  
 Db |||||  
 QY 2108 CATCTGAATGTACAGAGAAAGGCTCTCGATCACATCTATAGCAATTTCTGGAGCGCG 2167  
 Db |||||  
 QY 2462 CTGAAGGAATGAATGGGAGTGGCCAAACCCAGGAGCAATGCCCATTTGTGTGAGA 2521  
 Db |||||  
 QY 2168 GTGAAGGAATGTACAGAGAGAGTTCTGAAGCCCG---ACGACGTCCTTCTGGAGCGCG 2224  
 Db |||||  
 QY 2522 AGACTGCTGTTCAACACTGCAAGAGTGTGAGCTGTTTATATGATGACAGAGATGGCTTT 2581  
 Db |||||  
 QY 2225 AAGATGCTTTTCGAGGAGCAAGAGTGTGATGATTTCAAGATGGAGATGGATTC 2284  
 Db |||||  
 QY 2582 GGTATCTCTGCAAGAGATGAAGACCATGTCTAGCCGAACTCTTTTGGATCTCTGTGGCG 2641  
 Db |||||  
 QY 2285 GGTGTTTCCAAATAGAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGGTGCA 2344  
 Db |||||  
 QY 2642 TAGCA-----TACTGATATATATATATATATATATATATATATATATATATAT 2688  
 Db |||||  
 QY 2345 CTGTAATCAAAATAGTTGCAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2404  
 Db |||||  
 QY 2689 AAAAAAAAAAAAAA 2705  
 Db |||||  
 QY 2405 AAAAAAAAAAAAAA 2421  
 Db |||||

RESULT 6  
 AAX87529

ID AAX87529 standard; cDNA; 2528 BP.

AC AAX87529;

XX

DT 08-OCT-1999 (first entry)

DE Grand fir E-alpha-bisabolene synthase cDNA.

XX E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;

KW gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour; ss.

XX Abies grandis.

XX Key Location/Qualifiers

FT CDS 1. .2454

FT /\*tag= a

FT replace (194,T)

FT /\*tag= b

FT /note= "T in clone AGI of AAX87504"

XX WO9937139-A1.

XX 29-JUL-1999.

XX

PF 21-JAN-1999; 99WO-US001300.

XX

XX 22-JAN-1998; 98US-0072204P.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;

XX WPI; 1999-458569/38.

XX P-PSDB; AAY06566.

XX Grand fir (*Abies grandis*) sesquiterpene synthase genes and protein products, useful for sesquiterpenoid production.

XX Example 11; Page 89-92; 136pp; English.

XX This is the nucleotide sequence of grand fir (*Abies grandis*) E-alpha-bisabolene-synthase cDNA. It codes for a wound-inducible enzyme (see AAY06566) that is capable of generating multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl diphosphate. The full-length cDNA was obtained by 5'RACE from a wound-induced grand fir stem cDNA library. It differs from the sequence of clone AGI, provided in AAX87504, by a single nucleotide at position 194 (C is AGI, T in the present sequence) that results in an amino acid change from Ala-65 in AGI to Val-65 in the protein of AAY06566. The present sequence is believed to be the correct sequence of grand fir E-alpha-bisabolene synthase cDNA. The invention provides gymnosperm sesquiterpene synthase nucleic acids and recombinant proteins, as well as expression vectors, and a method of imparting or enhancing production of a gymnosperm sesquiterpene synthase in a prokaryotic or eukaryotic host cell, especially in a plant, in order to impart, enhance or modify production of sesquiterpenoids, and hence to alter the flavour or fragrance of plant products, to improve disease resistance or to alter ecological interactions mediated by farnesyl diphosphate and its derivatives

XX SQ Sequence 2528 BP; 750 A; 497 C; 600 G; 681 T; 0 U; 0 Other;

Query Match 27.08; Score 730.2; DB 2; Length 2528;  
 Best Local Similarity 59.6%; Pred. No. 8.6e-187;

Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;

QY 434 AATCCATCTGCATATGATACAGCTTGGTGGCAGAAATTCGTCATTTGACGGCTCTGGT 493

Db |||||

QY 238 ACTCCATCTGTTATGACAGCATGGGTAGCGAGGGTCCCGCATTTGCTTGGTCTGCT 297

Db |||||

QY 494 GCACCCCAATTTCCCAACCGTTCAATGGATTCTGAACATCAACTGCCAGATGGCTCG 553

Db |||||

QY 298 CGCCCGCAATTTCCCAACCGTTCAATGGATTCTGAACATCAACTGCCAGATGGCTCG 357

Db |||||

QY 554 TGGGTGAGGAGTGCAATTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTCGCTC 613

Db |||||

QY 358 TGGGGAATTCAGTCCCACTTTCTGCTGCGACCGTCTTCTTGTGCTCTTCTTGTGTT 417

Db |||||

QY 614 CTCACCTCAAAATATGGAATAGGGCGACATTCAGTGCAGAAAGGGTGTGATTTGTG 673

Db |||||

QY 418 CTTGTGCTCTTAAATGGAACGTTGGGATCTGCAATAGCAGGGAATTAATTCATA 477

Db |||||

QY 674 AGAAAAACATGGAAGAAATGAAGGACGAAGCTGACAAATCACAG---GCCAAGTGGATTC 730

Db |||||

QY 478 AAGAGCAATCTGGAACCTAGTAAAGGATGAAACCGATCAAGATAGCTTGGTAAACAGACTTT 537

Db |||||

QY 731 GAGTCTGTTTCTCTGCAATGTTAGATGAACAAAAGCTTGGATTTGGATCTTCTTAT 790

Db |||||

QY 538 GAGATCATATTTCTCTCTGTTAAGAGAAGCTCAATCTCTGCGCTCTCGGACTTCCCTAC 597

Db |||||

QY 791 CACCTCCCTTTTCATCTCCCAAAATCCCAAAAGCGCCAGAAAAGCTTCAAAAGATTTCCC 850

Db |||||

QY 598 GACCTGCTTAT 657

Db |||||

QY 851 CTCATATGTTCTTCAATACCATCAGACGGCGTGTGCTCTACTCTCTGTGAGGGTTTGCAGAT 910

Db |||||

QY 658 AGGGAGGAATTTATGCGGTTCCGTCGCCATTTGTTGATTCTTTAGAGGAATACAAGAT 717

Db |||||

QY 911 GTGGTGACTGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTTAAGCTCC 970

Db |||||

D	b	718	ATAGTGAATGGGAACGAATTAATGAAGTTCAAAAGTCAGGATGGTCTTTCTTAAGCTCA	777	1795	TAT---ATGAAACTATGTTACCAAACTATTATGACATAGTTCCAGAGGTGGCTTGGGAG	1851
Q	y	971	CCTGCATCTACTGCTGTGCTTCTCATGACACTCAAAACAAACATGCTCCTCACTTTCTC	1030	2051	GGACTCAAGGAAACAAGCCGTGATGTCTGGGCTATCTTCGAAAAGTATGGAGGGCTTG	2110
D	b	778	CCTGCTTCTACTGCTGCTGCTTTTCTATGACACAGAGAGCGGAATGCCITGAATCTTG	837	1852	GCAGAGAAGGAACAAGGGCGTGATTTGGTCAGCTTTTCAGAAAGGATGGAGGATTAT	1911
Q	y	1031	AACTTCTGCTCAGCAAAATTTGGCGACTACGTTCTTGCCATATACCCACTTGATCTATT	1090	2111	CTGCGATCGTATATACCAAGAACCGAATGTCGGCAGCAAAAGTATGTGCCGACATTCAC	2170
D	b	838	AACAGTGTGATGATCAAGTTTGGAAATTTGTTCCCTGCTGTATCTCTGTGGATCTGCTG	897	1912	CTTCTGGGTTATTATGAAGAAGCTGAATGGTGTAGCTGCTGAGTATGTGCCCTACCTTGGAC	1971
Q	y	1091	GAAAGCCTCTGGGCTGTCGATACAGTTGAACGCTTGGGAATCGATTCGATTTCAAGAAA	1150	2171	GAATATGTGAAAATGCAAAAGTGTCCATAGCACTTCGGCAGCTGCTACTAAACTCAATC	2230
D	b	898	GAAAGCCTGTTGATCGTATAGTAATATGTAGCGCTTGGAACTATAGACACTTTGAANAAG	957	1972	GAGTACATAAAGAAATGGAATCACAATCTATCGCCCAAGTATACTTCTGTTGAGTGGAGTG	2031
Q	y	1151	GAAATCAAAAGAACTCTCGATTAACGTTTATAGSTACTGGGACGCCGAAGAGGGGTGGGA	1210	2231	TTTTTCA---CTGAGAGAATTACTTCTCGATTTACATTTTACAGCAAGTAGACTTCGGTCC	2287
D	b	958	GAAATCAAGGAAGCTCTTGATTTATGTTTACAGGCATTTGGACG---AAAGAGGAATTTGG	1014	2032	TTGATATGATGGGCAACTCTTTCGCAAGAGGCAATTAGAGAAAGTAGATTATCCAGGA	2091
Q	y	1211	TGGGCAAGATCAATCTTATCTGATGTCGATGACACTGCCATGGTCTTGAATCCCTG	1270	2288	AAATTTCTGC-----ATCTTGTGCTCTTTGACTGGAGCACTAACTCAATGACACCAAGACT	2341
D	b	1015	TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTTCGATGCTT	1074	2092	AGAGTGTCTTCACAGAGCTGAATAGCCTCATTTTCCGCGCTGGCGGATGACACGAGACA	2151
Q	y	1271	AGACTTCTATGGATACAACTATCTTCTCAGATGTTCTGGAGAAATTTTCAGAGACGAGAAAGGA	1330	2342	TACCAGGCCGAGAGAAAACCGTGGTGAATTTGGTTCAGCGGTACAGTGTCTACATGAGGAA	2401
D	b	1075	CGGCTGCATAGTACATATGATCTCTCAGCCATTTTTCAGCACTTTCAAGATGCCAATGG	1134	2152	TATAAGCTGAGAAGGCTCTGGAGAAATTTGGGTCAGCACTTTGAATGTTACATGAAGAC	2211
Q	y	1331	GACTTCTTTTGTTCGGGTCAAAAGCAAAATTTGTTGACCGATTAATCTTAACCTTTAT	1390	2402	AATCCGGAGTGCACAGAGGAAGAGCTCTAAATGATTTGTTATGTTATGTTATCATCGACAACCA	2461
D	b	1135	AAATTCATTTGCTCGACCGCTCAATTCMAACAAGATGTAGCAAGCATGCTGAATCTTTAT	1194	2212	CATCTGAAATGTACAGAGGAAGAGGCTCTCGATCATCATATAGCAATTTCTGGAGCCGGCG	2271
Q	y	1391	AGATGTTCAAGTATGTTTTCGGGAGAAAAGATAATGGAAGAGCTAAGACCTTCACT	1450	2462	CTGAAGGAATTTGAAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCAATTTGTGTGAGA	2521
D	b	1195	AGAGCTTCCAGCTCGCATTTCCCGGAGAAAACAATCTTGTATGAGCTTAAAGCTTGCT	1254	2272	GTGAAGGAACCTGCAAGAGAGTTTCTGAAGCCCG---ACGAGCTCCCATTCGCTGCAAG	2328
Q	y	1451	ACAAATCATCTCCAAATGCTCTTGCCAAAACAACGCAATTTGATTAAGTGGGCTGTCAAG	1510	2522	AGACTGCTGTTTCAACACTGCAAGAGTGTATGCAAGCTGTTTATATGTATACAGATGCTCTT	2581
D	b	1255	ACTAATATTTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAACAAACAA	1314	2329	AGATGCTTTTCGAGGAGACAAAGAGTGAAGTGTGATATTCAAGGATGGAGATGGATTTC	2388
Q	y	1511	AGGATCTTCTCGGAGGTTGGATGATCTATTAAGTATCCGTCGATAGAGATGATGCCA	1570	2582	GGTATCTCTGCAAAAGAGATGAAGAACCATGTCAGCCGAACTCTTTTCGATCCTGTGGCG	2641
D	b	1315	CAAAACCTGAGCCCAAGAGATCAAAATACGGCTGAAGACTTTTGGCATGCGAGTGTCCG	1374	2389	GGTGTTCCTCAAAATTAGAAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCCA	2448
Q	y	1571	AGATTGGAGCAAGAGTATACATAGAGCAATTTGGATCAAAATGATGCTGCTGGCGGAAG	1630	2642	TAGCA-----TACTGATATTATATATATATATATTTCAATTCATATCCAAATCCAAAAAAA	2688
D	b	1375	AGATGGAACAAAGAGATCTCAAGTGTATCGCCAGATTTATGACGCA TAGCAAAA	1434	2449	CTGTAATCAAAATAGTTGCAATAATTAATTTGAAATTAATCTCAACTATGTTTTCACAAAAAA	2508
Q	y	1631	ACTGTGTATAGATGCTATATGTAGCAACGAAATAATTTGAGCTGGCCAAATTTGGAC	1690	2689	AAAAAAAACCAAAAAA 2705	
D	b	1435	TGGGTTTACAAGCTACCTACGTGAACATGAAAGTTTTTTAGAGCTGGGAAATTAGAT	1494	2509	AAAAAAAAAAAAAAAAA 2525	
Q	y	1691	TTCAATATGTCGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTACGCTGGTGAGA	1750			
D	b	1495	TTCAACATATCCAGTCCATCCACCAAGAGAAATGAAGATGTACCAGCTGGTTAGA	1554			
Q	y	1751	GAATCGGGAATCAATGATCTTACATTCACCGCCAGCGGCTGTGGAAATGATTTCTCA	1810			
D	b	1555	GATTCGGGGTTGCCACTATTACCTTCGCTCGGAGAGGGCGCTGGAAATTTCTACTCTTA	1614			
Q	y	1811	GTGGCGTTAGTATGTTTGGCCAGCAATTCGCTGTTGTAGAAATTTGCTATGCCAGACT	1870			
D	b	1615	GTAGCGGGGGACCTATGAACCCAGTATGCCAAATGCGAGTTCTCTTTTACAAAAGTG	1674			
Q	y	1871	TCTTGCTCGCAGTATTTCTAGACGATCTTTTACGACACCCACCGATCTCTGGATGATCTT	1930			
D	b	1675	GCATGCTTGCAGACTGTTCTGACGATATGATGACACTTATGGAACCCCTAGATGAATTG	1734			
Q	y	1931	AAATGTTCTCTGAAGCGGTCGAGATGGAATCTCTGTGCTGATAGGTTTCGGGAT	1990			
D	b	1735	AAGCTATTCACTGAGGCTGTGAGAAGATGGGACCTCTCTTTTACAGAAAACCTTCCAGAC	1794			
Q	y	1991	AATCAGTTGAAAGTTTGTCTTCTAGGCTGTATCAACACAGTGAATGGATTTGGAAAAGAT	2050			

RESULT 7  
AX87531  
ID AX87531 standard; DNA; 2528 BP.  
XX  
AC AX87531;  
XX  
DT 08-OCT-1999 (first entry)  
XX  
DE E-alpha-bisabolene synthase computer-generated nucleic acid sequence.  
XX  
KW E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;  
KW gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour; ss.  
XX Synthetic.  
OS Abies grandis.  
XX  
FH Key 1.2454 Location/Qualifiers  
FT CDS /\*tag= a  
XX  
FN W09937139-A1.



PD	29-JUL-1999.	Db	718	ATAGTTGAATGGGAAAGAAATTAATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTAAGCTCA	777
XX					
PF	21-JAN-1999; 99WO-US001300.	Qy	971	CTTGCATCTACTGCTTGTGTCTTTCATGCACTCAAAACAAACGATGCTCCACTTTCTC	1030
XX					
PR	22-JAN-1998; 98US-0072204P.	Db	778	CTGCTTCTACTGCTGCGTTTTTATGCAACAGGAGACGCGAAATGCTTTGAATCTTG	837
XX					
PA	(UNIW ) UNIV WASHINGTON STATE RES FOUND.	Qy	1031	AACCTTGTCTCAGCAAAATTTGGCGACTAGCTTCTTGGCAATTTACCACTTGTATTTT	1090
XX					
PI	Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;	Db	838	AACAGTGTGATGATCAAGTTTGGAAATTTTGTCTCTGCTGCTATCTCTGAGATCTGCTG	897
XX					
DR	WPI; 1999-458569/38.	Qy	1091	GAACGCTCTGGGCTGTTCGATCAGTTTGAACGCTTGGGAATTCGATCCCTATTTTCAAGAA	1150
DR	P-PSDB; AAY06568.	Db	898	GAACGCTGTGTGATCGTAGATAATTTGTAGCGCTTGGATCTTATAGACACTTTGAAAG	957
XX					
PT	Grand fir ( <i>Abies grandis</i> ) sesquiterpene synthase genes and protein	Qy	1151	GAATCAAAAGAAATCTCTGGAATAGCTTTATAGTACTTGGGACGCGGAAAGAGCGGTGGGA	1210
PT	products, useful for sesquiterpenoid production.	Db	958	GAATCAAGGAGAGCTCTTGATTTTACAGGCAATTTGGAACG---AAAGAGGAATTTGGG	1014
XX					
PS	Example 11; Page 101-104; 136pp; English.				
XX					
CC	This is the sequence of a computer-generated nucleic acid coding for a	Qy	1211	TGGGCAAGATGCAATCTTATCTCTGATGTCGATGACACTGCCCATGGTCTTAGAATCTCTG	1270
CC	grand fir ( <i>Abies grandis</i> ) E-alpha-bisabolene synthase polypeptide (see	Db	1015	TGGGGCAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTTCGATTCCTT	1074
CC	AAY06568) that has conservative amino acid substitutions relative to the	Qy	1271	AGACTTTCATGGAATACAAATGTATCTTCAGATGTTTCTGAGAGTCTTCAGAGACGAGAAAGGA	1330
CC	native E-alpha-bisabolene synthase sequence given in AAY06562. E-alpha-	Db	1075	CGCTGCTAGTAGTACATGATGTAATCTTCAGCAATTTTTCAGCACTTCAAGATGCCAATGGG	1134
CC	bisabolene synthase is a wound- inducible enzyme capable of generating				
CC	multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl	Qy	1331	GACTTCTTTTGTCTTTCGCGTCAACCGCAAAATTTGGTGTGACCGATTAATCTTAAACCTTTAT	1390
CC	diphosphate. The invention provides gymnosperm sesquiterpene synthase	Db	1135	AAATTCATTTGCTCGACCGGCTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1194
CC	nucleic acids and recombinant proteins, as well as expression vectors,				
CC	and a method of imparting or enhancing production of a gymnosperm	Qy	1391	AGATGTTTCAAGTATGTTTTCGGGAGAAAAGATTAATGGAAGAGCTTAAGACCTTCACCT	1450
CC	sesquiterpene synthase in a prokaryotic or eukaryotic host cell,	Db	1195	AGAGCTTCCAGCTCGCAATTTTCCCGGAGAAAACAATCTTGTGATGAAGCTTAAAGCTTCGCT	1254
CC	especially in a plant, in order to impart, enhance or modify production	Qy	1451	ACAAATCATCTCCAAAATGCTCTTTCGCAAAAACAAACGCAATTTGATAGTGGCTGTCAAG	1510
CC	of sesquiterpenoids, and hence to alter the flavour or fragrance of plant	Db	1255	ACTAAATATTTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGATGAGAAACAAACA	1314
CC	products, to improve disease resistance or to alter ecological	Qy	1511	AAGGATCTTCTCGGAGAGGTGGAGTATGCTATATAAGTATCTCGGTGGCATAGAAAGTATGCCA	1570
CC	interactions mediated by farnesyl diphosphate and its derivatives	Db	1315	CAAAACCTGAGCCACAGAGATCAAAATACGCGCTCAGAGACTTCTTGGCATGCCAGTGTTCG	1374
XX					
SQ	Sequence 2528 BP; 750 A; 498 C; 599 G; 681 T; 0 U; 0 Other;	Qy	1571	AGATTGGAGGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG	1630
		Db	1375	AGAGTGGAGCAAGAGATACCTGCTCAAGTGTATCGCCAGATTTATGCCACATAGCAAAA	1434
		Qy	1631	ACTGTGTATAAGATGCTATATGTGAGCAACGAAAATATTTGGAGCTTGGCCAAATTTGGAC	1690
		Db	1435	TGCGTTTACAGCTACCTACGTGAACATGAAAAGTTTTTTAGAGCTGGGAAAATTAGAT	1494
		Qy	1691	TTCAATATGCTGAGGCTTTTACACCAAAAGGAGACTCAACACATTTGTACGTGGTGAGA	1750
		Db	1495	TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAAGAAATGTTACCACTGGTGTAGA	1554
		Qy	1751	GAATCGGGATTCAATGATCTTACATTCACCCGCGAGCGGCTGTGGAAATGTTATTTCTCA	1810
		Db	1555	GATTTCGGGTTGCGCACTATTTCACCTTCGCTCGGAGAGGCGCTGGAAATTTCTACTCTTA	1614
		Qy	1811	GTGCGGTTAGTATGTTTGAAGCCAGAAATTCGCTGCTTGTAGAATTCCTATGCCAAGACT	1870
		Db	1615	GTAGCGGGGGAGCTTATGAACCCCGATGATGCCAAATGCAAGGTTTCTCTTTTACAAAAGTG	1674
		Qy	1871	TCCTTGCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCACCGATCTCTGGATGATCTT	1930
		Db	1675	GCATGCTTGCAGACTGTTCTTGGACGATATGATAGACACTTTATGGAACCTTAGATGAATG	1734
		Qy	1931	AAATGTTTCTCTGAAGCGGTCCGAAAGATGGGATATCTCTGTGCTGGATAGCGGTTCCGGAT	1990
		Db	1735	AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCTCTTTACAGAAAAACCTTCCAGAC	1794
		Qy	1991	AATCAGTTGAAAGTGTGCTTCTTAGGCTGTACACACAGTGAATGGATTTGGAAGAAGAT	2050
		Db	1795	TAT---ATGAAACTATGTTTACCAAATCTATTATGACATAGTTTACGAGGTGGCTTGGAG	1851

PD	29-JUL-1999.	Db	718	ATAGTTGAATGGGAAAGAAATTAATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTAAGCTCA	777
XX					
PF	21-JAN-1999; 99WO-US001300.	Qy	971	CTTGCATCTACTGCTTGTGTCTTTCATGCACTCAAAACAAACGATGCTCCACTTTCTC	1030
XX					
PR	22-JAN-1998; 98US-0072204P.	Db	778	CTGCTTCTACTGCTGCGTTTTTATGCAACAGGAGACGCGAAATGCTTTGAATCTTG	837
XX					
PA	(UNIW ) UNIV WASHINGTON STATE RES FOUND.	Qy	1031	AACCTTGTCTCAGCAAAATTTGGCGACTAGCTTCTTGGCAATTTACCACTTGTATTTT	1090
XX					
PI	Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;	Db	838	AACAGTGTGATGATCAAGTTTGGAAATTTTGTCTCTGCTGCTATCTCTGAGATCTGCTG	897
XX					
DR	WPI; 1999-458569/38.	Qy	1091	GAACGCTCTGGGCTGTTCGATCAGTTTGAACGCTTGGGAATTCGATCCCTATTTTCAAGAA	1150
DR	P-PSDB; AAY06568.	Db	898	GAACGCTGTGTGATCGTAGATAATTTGTAGCGCTTGGATCTTATAGACACTTTGAAAG	957
XX					
PT	Grand fir ( <i>Abies grandis</i> ) sesquiterpene synthase genes and protein	Qy	1151	GAATCAAAAGAAATCTCTGGAATAGCTTTATAGTACTTGGGACGCGGAAAGAGCGGTGGGA	1210
PT	products, useful for sesquiterpenoid production.	Db	958	GAATCAAGGAGAGCTCTTGATTTTACAGGCAATTTGGAACG---AAAGAGGAATTTGGG	1014
XX					
PS	Example 11; Page 101-104; 136pp; English.				
XX					
CC	This is the sequence of a computer-generated nucleic acid coding for a	Qy	1211	TGGGCAAGATGCAATCTTATCTCTGATGTCGATGACACTGCCCATGGTCTTAGAATCTCTG	1270
CC	grand fir ( <i>Abies grandis</i> ) E-alpha-bisabolene synthase polypeptide (see	Db	1015	TGGGGCAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTTCGATTCCTT	1074
CC	AAY06568) that has conservative amino acid substitutions relative to the	Qy	1271	AGACTTTCATGGAATACAAATGTATCTTCAGATGTTTCTGAGAGTCTTCAGAGACGAGAAAGGA	1330
CC	native E-alpha-bisabolene synthase sequence given in AAY06562. E-alpha-	Db	1075	CGCTGCTAGTAGTACATGATGTAATCTTCAGCAATTTTTCAGCACTTCAAGATGCCAATGGG	1134
CC	bisabolene synthase is a wound- inducible enzyme capable of generating				
CC	multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl	Qy	1331	GACTTCTTTTGTCTTTCGCGTCAACCGCAAAATTTGGTGTGACCGATTAATCTTAAACCTTTAT	1390
CC	diphosphate. The invention provides gymnosperm sesquiterpene synthase	Db	1135	AAATTCATTTGCTCGACCGGCTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1194
CC	nucleic acids and recombinant proteins, as well as expression vectors,	Qy	1391	AGATGTTTCAAGTATGTTTTCGGGAGAAAAGATTAATGGAAGAGCTTAAGACCTTCACCT	1450
CC	and a method of imparting or enhancing production of a gymnosperm	Db	1195	AGAGCTTCCAGCTCGCAATTTTCCCGGAGAAAACAATCTTGTGATGAAGCTTAAAGCTTCGCT	1254
CC	sesquiterpene synthase in a prokaryotic or eukaryotic host cell,	Qy	1451	ACAAATCATCTCCAAAATGCTCTTTCGCAAAAACAAACGCAATTTGATAGTGGCTGTCAAG	1510
CC	especially in a plant, in order to impart, enhance or modify production	Db	1255	ACTAAATATTTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGATGAGAAACAAACA	1314
CC	of sesquiterpenoids, and hence to alter the flavour or fragrance of plant	Qy	1511	AAGGATCTTCTCGGAGAGGTGGAGTATGCTATATAAGTATCTCGGTGGCATAGAAAGTATGCCA	1570
CC	products, to improve disease resistance or to alter ecological	Db	1315	CAAAACCTGAGCCACAGAGATCAAAATACGCGCTCAGAGACTTCTTGGCATGCCAGTGTTCG	1374
XX					
SQ	Sequence 2528 BP; 750 A; 498 C; 599 G; 681 T; 0 U; 0 Other;	Qy	1571	AGATTGGAGGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG	1630
		Db	1375	AGAGTGGAGCAAGAGATACCTGCTCAAGTGTATCGCCAGATTTATGCCACATAGCAAAA	1434
		Qy	1631	ACTGTGTATAAGATGCTATATGTGAGCAACGAAAATATTTGGAGCTTGGCCAAATTTGGAC	1690
		Db	1435	TGCGTTTACAGCTACCTACGTGAACATGAAAAGTTTTTTAGAGCTGGGAAAATTAGAT	1494
		Qy	1691	TTCAATATGCTGAGGCTTTTACACCAAAAGGAGACTCAACACATTTGTACGTGGTGAGA	1750
		Db	1495	TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAAGAAATGTTACCACTGGTGTAGA	1554
		Qy	1751	GAATCGGGATTCAATGATCTTACATTCACCCGCGAGCGGCTGTGGAAATGTTATTTCTCA	1810
		Db	1555	GATTTCGGGTTGCGCACTATTTCACCTTCGCTCGGAGAGGCGCTGGAAATTTCTACTCTTA	1614
		Qy	1811	GTGCGGTTAGTATGTTTGAAGCCAGAAATTCGCTGCTTGTAGAATTCCTATGCCAAGACT	1870
		Db	1615	GTAGCGGGGGAGCTTATGAACCCCGATGATGCCAAATGCAAGGTTTCTCTTTTACAAAAGTG	1674
		Qy	1871	TCCTTGCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCACCGATCTCTGGATGATCTT	1930
		Db	1675	GCATGCTTGCAGACTGTTCTTGGACGATATGATAGACACTTTATGGAACCTTAGATGAATG	1734
		Qy	1931	AAATGTTTCTCTGAAGCGGTCCGAAAGATGGGATATCTCTGTGCTGGATAGCGGTTCCGGAT	1990
		Db	1735	AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCTCTTTACAGAAAAACCTTCCAGAC	1794
		Qy	1991	AATCAGTTGAAAGTGTGCTTCTTAGGCTGTACACACAGTGAATGGATTTGGAAGAAGAT	2050
		Db	1795	TAT---ATGAAACTATGTTTACCAAATCTATTATGACATAGTTTACGAGGTGGCTTGGAG	1851



Db	702	AGGGAGGAAATTATGCGGTTCCGTGCGCATGTGTGTAATCTTTAGAGGGAAATACAAGAT	761
Qy	911	GTGTGGACTGGCAAGAGATCAAAATCTTTCAATCAAGAGACGAGATCAATTTTTAAAGCTCC	970
Db	762	ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTTAAGCTCA	821
Qy	971	CTGTGATCTACTGCTTGTGTCTTCATGCACTCACTCAAAACAAACGATGCCCTCCACTTTCTC	1030
Db	822	CCTGCTTCTACTGCTCGGTTTTCATGCACACAGGAGACGCGAAATGCCTTGAATCTCTTG	881
Qy	1031	AAC TTCGTGCTCAGCAAAATTTGGCGACTACGTCTCTTGGCAATTTACCCACTTTGATCTATTT	1090
Db	882	AACAGTGTGATCAAGTTTGGAAAATTTGTGTCCTGCTGTATCTGTGGATCTGCTG	941
Qy	1091	GAA CGCCTCTGGGCTGTCGATACAGTTGAA CGCTTTGGAAATCGATCGCTATTTCAAGAAA	1150
Db	942	GAA CGCCTGTGATCGTAGATAAATTTGTAGCCCTTGGAACTATAGACACTTTGAAAG	1001
Qy	1151	GAAATCAAAGAAATCTCTGGATTACGTTTTATAGTTACTGGGACGCCGAAAGAGCGGTGGGA	1210
Db	1002	CAAAATCAAGGAAGCTCTTGATTATGTTTACAGGCATTTGGAACG---AAAGAGGAATTTGG	1058
Qy	1211	TGGCGAGATGCAATCCTATTTCTGTGATCGATGACACTGCCATGGGTCTTTAGAACTCTG	1270
Db	1059	TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTTCGATGCTTT	1118
Qy	1271	AGACTTATGATACAATGTATCTTCAGATGTTCTGGAGAAATTTCAAGACGAGAAAGGA	1330
Db	1119	CGGCTGCATAGGTACAATGATATCTCAGCCATTTTGCACAACTTCAAAGATGCCATGGG	1178
Qy	1331	GACTTCTTTTGGCTTTGGCGGTCAAA CGCAATTTGGTGTGACCGGATAATCTTAACCTTTAT	1390
Db	1179	AAATTCATTTGCTCGACCGGTCAATTCAA CAAAGATGTAGCAAGCATGCTGAAATCTTTAT	1238
Qy	1391	AGATGTTCACAAGTATGTTTCCGGGAGAAAGATAATGGAAGACTAGACCTTCACT	1450
Db	1239	AGAGCTTCCCAGCTGCGATTTCCCGGAGAAAACAATCTTGATGAAGCTTAAAGCTTCGCT	1298
Qy	1451	ACAAATCATCTC AAAATGCTCTTTGCCAAAAA CAA CGCATTTGATAGTCGGCTGTCAAG	1510
Db	1299	ACTAAATATTTGAGAGAGCTCTTTGAGAAAGTGCAGACTTCCAGTGCATGGAACACAA	1358
Qy	1511	AAGGATCTTCTCGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAGTATGCCA	1570
Db	1359	CAAAACCTGAGCCAAGAGATCAAAATACGCGCTGAAGACTTCTTGGCATGCCAGTGTCCG	1418
Qy	1571	AGATTTGGAGGCAAGATTACATAGAGCAATTTGGATCAAATGATGTCTGGCTGGGGAG	1630
Db	1419	AGAGTGGGAAGCAAGAGATACTGTCAAAGTGTATGCCCCAGATTATGCACGCATAGCAAAA	1478
Qy	1631	ACTGTATTAAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGCCAAATGGAC	1690
Db	1479	TGGTTTTACAAGCTACCTTACGTGAA CAA TGAAGTTTTTTAGAGCTTGGGAAAATTTAGAT	1538
Qy	1691	TTCAATATGTCAGGCGCTTACCA CAAAAGGAGACTCAACACATTTGCTAGCTGGTGAGA	1750
Db	1539	TTCAACATTATCCAGTCCATCCA CCAAGAGAAATGAAGAATGTTACCAGCTGCTTTTGA	1598
Qy	1751	GAATCGGGATTCATATGATCTTTACATTCACCGCCAGCGGCTGTGGAAATGTAATTTCTCA	1810
Db	1599	GATTCGGGGTTGCCACTATTCACCTTTCGCTCGGAGAGCGGCTGGAAATTTCTACTTCTTA	1658
Qy	1811	GTGCGGTTACTATGTTTGAAGCCAGAAATTCGCTCTTGATAGATTGCTCATGCCAAGACT	1870
Db	1659	GTAGGCGGGGACCTATGAACCCAGTATGCCAAATTCGAGGTTCTCTTTTACAAGAGT	1718
Qy	1871	TC TTGCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCACCGGATCTCTGGATGATCTT	1930
Db	1719	GCATGCTTGACAGCTGTTCTTCGACGATATGTA TGACACTTATGGAACCTTAGATGAATTG	1778
Qy	1931	AAATTTGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGTGGAATGAGGTTTCGGAT	1990
Db	1779	AAGCTATTTCACTGAGGCTGTGAGAAGATGGGACCTCTCTTTTACAGAAAACTTTTCAGAC	1838

QY	1991	AATCAGTTTGAAGTTTCTCTTAGGGCTGTACAAACAGTGAATGAATTTGGAAAAGAT	2050
DB	1839	TAT---ATGAAACTATGTATTACCAAACTATTTATGACATAGTTTACGAGGTGGCTTGGGAG	1895
QY	2051	GGACTCAAGGAAACAAGCCGTGTAGTCTGGCTTATCTTCCGAAAGTATGGGAGGGCTTG	2110
DB	1896	GCAGAAAGGAAACAGGGCGGTGAATTTGGTCAGCTTTTTCAGAAAGGATGGAGGATAT	1955
QY	2111	CTCGCATCGTATATCAAAAGAAGCCGAATTGGTCGGCAGCAAAAGTATGTGCCGACATTCACAC	2170
DB	1956	CTTCTGGGTATTATTAAGAAGCTGCTAGTCTGAGTATGTGCGCTACCTTGGAC	2015
QY	2171	GAATATGTGGAATAATGCCAAAGTGTCCATAGACCTTCGGCAGTCGTCGTAATAACTCAATC	2230
DB	2016	GAGTACATAAAGAAATGGAATCACATCTATTCGGCCAACTGATCTTCTGTGTGAGTGGAGTG	2075
QY	2231	TTTTTTTCA--CTCGAGAAATTACTTCTTCGTATTACATTTTTTACAGCAAGTAGACCTTCGGTCC	2287
DB	2076	TTGATAATGGAATGGCAACTCCCTTTCCGCAAGAGGCATTAGAGAAAGTAGAATTATTCAGGA	2135
QY	2288	AAATTTCTGC-----ATCTTGTGTCTTTTGACTGGACGACTAATCAATGACCAACAAGACT	2341
DB	2136	AGACGTGTCTTCACAGAGCTGAATAGCCTCATTTCCCGCCTGGCGGATGACACGAAGACA	2195
QY	2342	TACAGGCCGAGAGAGAAACCGTGTGTGAATTTGGTTTTCCAGCGTACAGTGCCTACATGAGGGAA	2401
DB	2196	TATAAAGCTCAGAAAGGCTCGTGGAGAAATTTGGCGTCCAGCANTGAATGTTTACATGAAAGAC	2255
QY	2402	AATCCGAGTGCACAGAGGAAGAGCTCTAAGTCATGTTTTATGGTATCATCGCAACAGCA	2461
DB	2256	CATCTCGAATGTACAGAGGAAGAGGCTCTCGATCACATCTATAGCATCTTGGAGCCGGCG	2315
QY	2462	CTGAAGGAAATTGAATTGGGAGTTTGGCCAAACCCAGCCGAGCAATGCCCAATTTGTGTGAGA	2521
DB	2316	GTGAAGGAACTGACAAGAGAGTTTCTGAAAGCCG---ACGACGTCCCAATTCGCCGTGCAAG	2372
QY	2522	AGACTGTCTTTCAACACTGCAAGAGTGAATGCACTGTGTATATGTAACAGAGATGGCTTT	2581
DB	2373	AAGATGCTTTTTCGAGGAGACAAGAGTGACCAATGGTGATATTCGAAGGATGGAGATTC	2432
QY	2582	GGTATCTTCACAAAGAGATGAAGACCAATGTGAGCCGAACTCTTTTCGATCTGTGCGCG	2641
DB	2433	GGTGTTTCCAAATTTAGAAGTCAAAGATCATTAAGAGAGTGTCTCATTTGAACCGCTGCCA	2492
QY	2642	TAGCA-----TACTGATATATATATAATTAATTCATATTCAATTCCAATCAAAAAAA	2688
DB	2493	CTGTAAATCAAAATAGTTGCAATAATAATTGAATATCTCAACTATGTTTCACAAAAAA	2552
QY	2689	AAAAAAAAAAAAAAAAAAAAA	2705
DB	2553	AAAAAAAAAAAAAAAAAAAAA	2569

RESULT 9  
AAT97447

AA19/447  
ID AAT97447 standard; cDNA; 2700 BP.

AC AAT97447;

XX DT 30-APR-19

XX  
DE  
RECEIVED

**XX**

KW FACILITIC yew; taxadiene synthase; taxol biosynthesis; probe; immunocassay;  
KW SS.

XX OS Taxing and

XX Vol:

FT	CDS	Location, quantity
FT	CDS	21. .2610

### LE

XX WO9738571-A1.  
XX 23-OCT-1997.  
XX 15-APR-1997; 97WO-US006320.  
XX 15-APR-1996; 96US-0015993P.  
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX Wildung MR, Croteau RB;  
XX WPI; 1997-526123/48.  
XX P-PSDB; AAW31655.  
XX  
XX DNA encoding Pacific yew taxadiene synthase - used to increase taxol biosynthesis.  
XX  
XX Claim 10; Fig 2; 30pp; English.  
XX  
XX This sequence encodes a pacific yew taxadiene synthase (TS). The sequence can be used to increase taxol biosynthesis into taxoids in transformed cells. The nucleic acid sequences can be modified to produce proteins with improved efficiency, and to redirect the targeting of the expressed polypeptide. The modifications can be to alter catalytic efficiency, to modify substrate utilisation, and to alter product outcome. The antibodies specific for the mature native pacific yew TS polypeptide can be used to purify TS polypeptides, cloning TS homologues from Pacific yew or other plant species, and as probes for blots and immunoassays  
XX  
XX Sequence 2700 BP; 819 A; 541 C; 616 G; 724 T; 0 U; 0 Other;  
XX  
Query Match 23.7%; Score 642.2; DB 2; Length 2700;  
Best Local Similarity 57.7%; Pred. No. 6.1e-163;  
Matches 1302; Conservative 0; Mismatches 913; Indels 42; Gaps 7;  
384 TGGTGAAGGAAATCCAGTGCATGTTTTCAGTCCATGGCGGACGCTGAAACGAATCCATCTG 443  
377 TGGTTGTGAATTAAGATATGTTCAATGCGCTCGGACGCGAGATATCAGTCCGCTG 436  
444 CATATGATACAGCTTGGGTGGCAAGAAATCCGTCAT------TGACGGCTCTGTGTCAC 497  
437 CATACACACTGCGTGGGTGGCGAGCTGGCGACCAATTCCTCTGATGATCTTGAGAAGC 496  
498 CCCAATTTCCCAAAAGCTTCAATGATTTCTGAACATCAACTGCCAGATGCTGCTGGG 557  
497 CACGGTTTCTCAGGCCCTCAACTGGGTTTTCAACCAACCAAGCTCCAGGATGGATCGTGG 556  
558 GTGAGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCCTCTCA 617  
557 GTATCGAATCGCACTTTAGTTTATGCGATCGATTGCTTAAACAGCAATTCGTATTGCG 616  
618 CTCTCAAAATATGGAATAGGGCGACATTCAGTGCAGAAAGGGTTGAGTTGTGAGAA 677  
617 CCCTCTCGGTTTGGAAAACAGGGGCACAGCCCAAGTACAAAGGTGCTGAGTTTATGCGAG 676  
678 AACACATGGAAGAAATGAAGGACGAGCTGCAACATCACAGGCCCAAGTGGATTTCAGGTGCG 737  
677 AGAATCTAAGATTACTCAATGAGGAGATGAGTTGTCCTCCG-----GATTCCAAATAA 730  
738 TGTTTCTCGCAATTTAGATGAAGCAAAAAGCTTGGGATTTGGATTTCTTATCACCTCC 797  
731 TCTTTCTGCTGCTGTCAAAAGGCAAAAGCGTTGGGGATCAATCTTCTCTAGCATCTTC 790  
798 CTTTCATCTCCCAATCCACCAAAAGGCCAGAAAGCTTCAAAAGATTTCCCTCAATG 857  
791 CAITTTATCAAAATATTTGTCGACAAACACGGGAAGCCAGGCTTACAGATGTTTCTGCG 847  
858 TTCTTCAATACCATCAGACGGCGTGTCTACTCTCTGAGGGGTTTGCAGAGATGTGGTGG 917  
848 CAGCAGCAATATTTCCAGCCCAACATGTTGAATGGCTTGGAAAGGTCTCGAGGAAGTTATTG 907

QY 918 ACTGCGAAGAGATCAAAATCTTCAATCAAGAGAGCGGATCAATTTTAAAGTCCCTCGAT 977  
DB 908 ACTGGACAAGATTATGAGGTTTCAAAGTAAAGATGATCTTTCTGAGCTCCCTGCTGCT 967  
QY 978 CTACTGCTTGTCTTTCATGACACTCAAAACAAACGATGCTCCACATTTCTCAACTTCG 1037  
DB 968 CCCTGCTGTTGTTATCTGATGAATACAGGGGACGAAAATGTTTCTCTTCTCAACAATC 1027  
QY 1038 TGCTCAGCAAAATTTGGCGACTACGTTCTCTGCGCAATACCCACTTGTATCTTATTGACGCC 1097  
DB 1028 TGCTCGACAAATTCGGCGCTGCGCTGCTGATGATTTCCATCGATCTGCTGGAACGCC 1087  
QY 1098 TCTGGGCTGTCTGATACAGTTGAAACGCTTGGGAATCGATCGCTATTTCAAGAAGAAATCA 1157  
DB 1088 TTTGCTGCTGTTGATAAATTCAGCATCTCGGAATCGGTGCGCATTTCAACAAAGAAATCA 1147  
QY 1158 AAGATCTCTGGATTACGTTTATAGGTACTGGAGCGCGAAGAGCGGTGGGATGGGCAA 1217  
DB 1148 AAGGAGCTCTTGTATGTTCTACAGACATTTGG---AGTGAAGGGGCGATCGGTTGGGCA 1204  
QY 1218 GATGCAATCTTATCTCTGATGTGATGACACTGCGCATCGGTCTTGAATCTCTGAGACTTC 1277  
DB 1205 GAGACAGCTTGTTCAGATCTCAACACACAGCCCTCGGCTGCGAACTCTTCGCAATGC 1264  
QY 1278 ATGGATACAAATGATCTTTCAGATGTTCTGGAAATTTCAAGACGAGAAAGAGACTTCT 1337  
DB 1265 ACGGATACAAATGTTTCTTTCAGACGTTTGAATTAATTTCAAGATGAAAACGGCGGTTCT 1324  
QY 1338 TTTGCTTTGCGGCTCAAAACGCAATTTGGTGTGACCGATAATCTTAACCTTTTATAGATT 1397  
DB 1325 TCTCTCTGCGGGCCAAACCCATGTGCAATTTGAAGAGCGGTGTTGAATCTTTTCAGAGCTT 1384  
QY 1398 CACAAAGTATGTTTTCGGGAGAAAAGATAATGGAAGAGCTTAAGACTTCACTACAAATC 1457  
DB 1385 CGACCTTGCATTTCTGACGAAAGAGCTATGAGAGCTAGTAAATTTGAGAACCAT 1444  
QY 1458 ATCTCAAAATGCTCTTCCCAAAACACAGCATTTGATAAGTGGGTGTCAAGAGGATC 1517  
DB 1445 ATCTTAGAGAGCATTTGCAACGAAATCTCA-----ACCAATACAAAC 1489  
QY 1518 TTCTCTGGAGGTGAGTATGCTATAAAGTATCCGTGGCATAGAAATATGCCAAGATTGG 1577  
DB 1490 TATTCAAGAGATTGATGAGTACGTTGGAGTACCTTTGGCAGATGATATCCCAAGCTAG 1549  
QY 1578 AGGCAAGAGTTTACATAGAGCAATTTGGATCAAAATGATGTTGGCTGGGAGAGACTGTGT 1637  
DB 1550 AAGCCAGAAGTTATATGATTCATATGACGCAATTTATGATGGCAGAGAGAGACTCTAT 1609  
QY 1638 ATAGATGCTATATGTGACCAACGAAAATATTTGGAGCTGGCCAAATTTGGACTTCAATA 1697  
DB 1610 ATAGAAATGCCATCTTTGAGTAATTTCAAATGTTTAGAATTTGGCAAAATTTGGACTTCAATA 1669  
QY 1698 TGGTGAGCGCTTTACACCAAAAGGAGACTCAACACATTTGTAGCTGGTGAGAGAAATCGG 1757  
DB 1670 TCGTACAAATCTTTGCATCAAGAGGAGTTGAAAGCTTCTAAACAGATGTTGGAAGAACCG 1729  
QY 1758 GATTCAATGATCTTACATTTCAACCGCGGCTGTGGAATGATTTCTCAGTGGCGG 1817  
DB 1730 GCATGCGAGATATAAATTTCACTCGACACCGAGTGGCGAGGTTTATTTTTCAGC-- 1787  
QY 1818 TTAGTATGTTTGGAGCCAGAAATTCGCTGTTGTAGAATTTGCTATGCGCAAGACTTCTTGCC 1877  
DB 1788 ----TACATTTGAAACCGAATATTTCTGCCCATAGATTTGCTTTCACAAAAATTTGGTTGT 1843  
QY 1878 TCGAGTTATTTAGACGATCTTTTACGACCCAGGATCTCTGGATGATCTTAAATGT 1937  
DB 1844 TACAAGTCTCTTTTGTATGATATGGCTGACATCTTTGCAACACTAGATGAATGAAAAGTT 1903  
QY 1938 TCTCTGAAGCGGTCCGAGATGGATATCTCTGCTGGATAGCGTTCCGGGATTAATCAGT 1997  
DB 1904 TCACTGAGGAGTAAAGAGATGGATACATCTTTTGTACATGAGATTCAGAGTGT---A 1960  
QY 1998 TGAAAGTTTCTTCTAGGCTGTCAACACAGTGAATGGAATTTGGAAAAGATGGACTCA 2057

Db 1961 TGCAAACTGCTTAAAGTTGGTTGTTCAATTAATGGAGAGTAATAATGATGTTGGTTA 2020  
Qy 2058 AGGAACAAGCCGCTGATGCTGGGCTATCTTCGAAAGATGATGGAGGGCTTCTCGCAT 2117  
Db 2021 AGGTACAAGGACGTGATGCTCGCTCACATAAGAAACCCCTGGAGTTGTACTTCAATT 2080  
Qy 2118 CGTATACCAAGAGCGGAATGCTGGCAGCAAAAGTATGCGGCACATCAACGAATATG 2177  
Db 2081 GTTATGTACAAGAAAGGAGTGGCTTGAAGCCGGTATATACCACTTTTGAAGTACT 2140  
Qy 2178 TGAATAATGCCAAAGTGTCCATAGCACTTGCAGACAGTCGTACTAAACTCAATCTTTTCA 2237  
Db 2141 TAAAGACTTATGCTATATCAGTAGGCTTGGACCGGTGACCTACCAACTACTACTAA 2200  
Qy 2238 CTGGAGAATTACTTCTGATTAATTTTACAGCAAGTAGACCTTCGGTCCAAATTTCTGC 2297  
Db 2201 TGGGTGAGCTTGTGAAGATGATGTTGTTGAGAAAGTGCACTATCCTCAATATGTTG 2260  
Qy 2298 ATCTTGCTGTTGACTGGACGACTAATCAATGACACCAAGACTTACCAGCCGAGAGAA 2357  
Db 2261 AGCTTGTATCTTGAGCTGGCGACTAACCAACGACACCAAAACATATCAGGCTGAAAGG 2320  
Qy 2358 ACCGTGTGAATTTGGTTTCCAGGTACAGTGTACATGAGGGAATCCGGAGTGCACAG 2417  
Db 2321 CTGAGGACAAAGCTCAGGCATGACATGCTATATGAAGGATATCCAGGAGCACTG 2380  
Qy 2418 AGGAAGAAGCTCTAAGTCTATGTTTATGATATCATCGACACGCACTGAAGGAATTGAAT 2477  
Db 2381 AGGAAGATGCCATTAAAGCACATATGCTGTTGTTGATGGGCTTGAAGAGCAAGCT 2440  
Qy 2478 GGGAGTTGGCAACCCAGCAGCAATGCCCCCATTTGTTGTGAGAGACTGCTCTTCAACA 2537  
Db 2441 TTGAATATTTCAAAACCATCAATGATATCCCAATGGGTTTCAAGTCTCTTATTTAAAC 2500  
Qy 2538 CTGCAAGAGTGTGACAGCTGTTTATATATGACAGAGATGCTTTGTTATCTCTGACAAAG 2597  
Db 2501 TTAGATTGTGTCCAAATCTTTTACAAAGTTTATAGATGGTACGGAATCCCAATGAGG 2560  
Qy 2598 AGATGAAGACCAATGTGACGCCGAACCTTTTTCGATCC 2634  
Db 2561 AGATTAAGGACTATATAAGAAAGTTTATATTGATCC 2597

## RESULT 10

AAA38931

ID AAA38931 standard; DNA; 2700 BP.

AC AAA38931;

XX 25-AUG-2000 (first entry)

DE Yew taxadiene synthase DNA sequence SEQ ID NO:43.

XX Synthase; protein co-ordinate data; active site; modification; terpenoid;  
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
KW defensive agent; pigment; antitumour; steroid hormone;  
KW signal transduction pathway; bile acid; affinity purification;  
KW photoreceptor; enzymatic synthesis; nutrient supplement;  
XX immunological reagent; ds.

OS Taxus brevifolia.

XX WO200017327-A2.

XX 30-MAR-2000.

XX 17-SEP-1999; 99WO-US021419.

XX 18-SEP-1998; 98US-0100953P.

XX 22-APR-1999; 99US-0130628P.

XX 23-AUG-1999; 99US-0150262P.

XX

PA (KENT ) UNIV KENTUCKY RES DEPT.

XX (SALK ) SALK INST BIOLOGICAL STUDIES.

PI Chappell J, Manna KR, Noel JP, Starks CM;

XX WPI: 2000-292839/25.

DR P-PSDB; AAY90852.

XX Novel terpene synthase enzymes, useful for producing terpene  
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
PT enzymes by specific amino acid alterations.

XX Disclosure; Page 408-412; 450pp; English.

XX The present invention describes an isolated terpene synthase (1)  
CC comprising a region with at least 20% identity to region 265-535 of a 548  
CC amino acid (aa) sequence (1a), given in AAY90831. (1) contains nine alpha  
CC -carbon atoms (alphaC) that have interatomic distances, between each  
CC other, within tabulated ranges, have a centre point (within a sphere of  
CC radius 2.3 Angstrom ) within tabulated ranges, and have an ordered  
CC arrangement of R groups (defining aa side chains), excluding specific  
CC tabulated arrangements (tables given in the specification). (1), and  
CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
CC components of signal transduction pathways, precursors of steroid  
CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
CC Some synthases with little or no catalytic activity (and nucleic acids  
CC encoding them) are used as controls in the analysis of products formed by  
CC enzymatic synthesis; as nutrient supplements; for affinity purification  
CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
CC monitoring expression of terpene synthase or inheritance of the gene in  
CC plant breeding programs. The new synthases may produce novel terpene  
CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
CC sequences used in the exemplification of the present invention

XX Sequence 2700 BP; 819 A; 541 C; 616 G; 724 T; 0 U; 0 Other;

Query Match 23.7%; Score 642.2; DB 3; Length 2700;  
Best Local Similarity 57.7%; Pred. No. 6.1e-163;  
Matches 1302; Conservative 0; Mismatches 913; Indels 42; Gaps 7;

Qy 384 TGGTGAAGAAATCCAGTCATGTTTCAGTCCATGGCGGACGGTGAACCAATCCATCTG 443

Db 377 TGGTTGTTGAAAATTAAGATATGTTCAATGCGCTCGGAGACGGAGATATCAGTCCGCTG 436

Qy 444 CATATGATACAGCTTGGGTGGCAAGAAATCCGTCAT-----TGACGGCTCTGGTGCAC 497

Db 437 CATACGACACTGCGTGGGTGGCGAGCTGGCGACCAATTCCTCTGATGATCTGAGAGC 496

Qy 498 CCCAATTTCCCAACCGCTTCAATGGATTCTGAACAATCAACTGCCAGATGGCTCGTGGG 557

Db 497 CAGGTTTCTCAGGCCCTCACTGGGTTTCAACAACCCAGCTCCAGGATGGATCGTGGG 556

Qy 558 GTGAGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTTCGCTCGCTCTCTCA 617

Db 557 GTATCAAAATCGCACTTTAGTTTATGCGATCGATTGCTTAAACAGCAACCAATTCGTGATCG 616

Qy 618 CTCTCAAAATATGAATAAGGCGACATTCAGTGCAGAAAGGGTTGAGTTTGTGAGAA 677

Db 617 CCCTCTCGGTTTGGAAAACAGGGGCAAGCCCAAGTACAAACAGGTGCTGATTTATTGCGAG 676

Qy 678 AACACATGGAAGAAATGAAGGACGAAGCTGACAAATCACAGGCCCAAGTGGATTTGAGGTGCG 737

Db 677 AGAATCTAAGATTACTCAATGAGAGAGATGAGTTGTCCCG-----GATTTCCAATAA 730

Qy 738 TGTTCCTCGCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGATCTTCTTATCACCTCC 797

Db 731 TCTTTCTGCTCTGCTGCAAAAAGGCAAAAGCGTTGGGGATCAATCTTCTCTTAGCATCTTC 790

Qy 798 CTTTCATCTCCCAATCCACCAAAAGGCCAGAAAAGCTTCAAAAGATTTCCCTCAATG 857

Db 791 CATTTATCAATATTTGTCGACACACCGGAGCCAGGCTTACAGATGTTTCTGC---GG 847  
Qy 858 TTCTTATCAACATCAGACGGGTTGCTCTACTCTCTGAGGGTTTGCAGATGTGGTG 917  
Db 848 CAGCAGACAATATTCAGCCAAACATGTTGAATGCGTTTGGAAAGGCTCGAGGAAGTTATTG 907  
Qy 918 ACTGCAAGAGATCAAAATCTTCAATCAAGAGCGGATCATTTTAAAGCTCCCTGTCAT 977  
Db 908 ACTGCAACAGATATGAGGTTTCAAGATTAAGATGGATCTTCTCTGAGCTCCCTGCT 967  
Qy 978 CTACTGCTGTGCTTTCATGCACTCAAAAACAAACGATGCTTCACTTCTCAACTTCG 1037  
Db 968 CCAGCTGCTGTACTGATGAATACAGGGAGCGAAATGTTTCACTTCTTCAACAATC 1027  
Qy 1038 TGCTCAGCAAAATTTGGGCACTAGTCTTCCATTCAGCAATCCAGCTTCACTTATTTGAACGCC 1097  
Db 1028 TGCTCAGCAAAATTTGGGCGCTGCGTCCCTGATGATTTCCATGATCTGCTGGAAACGCC 1087  
Qy 1098 TCTGGGCTGTCATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAAAGAAATCA 1157  
Db 1088 TTTCTGCTGTTGATTAACATTTAGCATCTCGGAATCGTCCGCAATTTCAACAAGAAATCA 1147  
Qy 1158 AAGAAATCTCTGGAATACGTTTATAGTACTGGGACCGGAAAGAGCGGTGGATGGGCA 1217  
Db 1148 AAGGAGCTCTGATTTATGCTACAGACATTTG---AGTGAAAGGGGCATCGGTTGGGCA 1204  
Qy 1218 GATGCAATCTTCTGATGTCGATGACATGCCATGGTCTTGAATCTCTGAGACTTC 1277  
Db 1205 GAGACAGCGCTTCTCGAGATCTCAACACACAGCCCTCGGCTGGCACTCTTCGATGC 1264  
Qy 1278 ATGGATACAAATGATCTTCAAGATGTTCTGGAGATTTTCAGACGAGAAAGAGACTTCT 1337  
Db 1265 ACGGATACAAATGTTCTTCAAGAGTTTGAATTAATTTCAAGATGAAGAAACGGCGGTTCT 1324  
Qy 1338 TTTGCTTCTCCGCTCAACCGAAAATTTGGTGTGACCGATTAATCTTAACTTTATAGTGT 1397  
Db 1325 TCTCTCTCGGGCCAAACCCATGTGCAATTTGAGAAGCGTGTGAAATCTTTTCAGAGCTT 1384  
Qy 1398 CACAGATATGTTTCCGGGAGAAAGATTAATGGAAGACTTAAGACTTCACTACAAATC 1457  
Db 1385 CGACCTTGCAATTTCTGACGAAGAGCTATGACGATCTAGAAATTTGAGAACCAT 1444  
Qy 1458 ATCTCAAAATGCTTCTGCAAAACCAACGCAATTTGATAGTGGGCTGTCAGAGAGATC 1517  
Db 1445 ATCTTAGAGAGGCACTTGCACAGAAATCTCA-----ACCAATACAAAAC 1489  
Qy 1518 TTTCTGGAGAGGTGGATATGCTATAAAGTATCCGTGGCATAGAGTATGCCAAGATTGG 1577  
Db 1490 TATTCAGAGAGATTGAGTACGTGGGTGGGTACCCCTTGGCACATGATATCCACGCTTAG 1549  
Qy 1578 AGGCAAGAGTTTACATAGAGCAATTTGGATCAAATGATGCTGGCTGGGGAAGACTGTGT 1637  
Db 1550 AAGCCAGAGTTATTTGATTCATATGACGCAATTTATGATGGCAGAGAGAGACTCTAT 1609  
Qy 1638 ATAGATGCTATATGTGAGCAACGAAATATTTTGGAGCTGGCCAAATTTGGACTTCAATA 1697  
Db 1610 ATAGAATGCCATCTTTGAGTAATTTCAAAATGTTTAGAATTTGGCAAAATTTGGACTTCAATA 1669  
Qy 1698 TGGTGGAGGCTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGTGGAGAGATCGG 1757  
Db 1670 TCGTACAATCTTTGGCATCAAGAGAGTTGAGCTTCTCAACAGATGTTGGAGAGATTCG 1729  
Qy 1758 GATTCATGATCTTACATTCACCCGCGCAGCGGCTGTGGAAATGTTATTTCTCAGTGGCGG 1817  
Db 1730 GCATGGCAGATATAAATTTTCACTCGACCGGAGTGGCGGAGGTTTATTTTTCATCAGC-- 1787  
Qy 1818 TTAGTATGTTTGAACCGAATTCGCTGTTGTAGATTTGCTATGCTATGCCAGACTTCTTGGC 1877  
Db 1788 ----TACATTTGAACCGGAATATTTCTGCCACTAGAATTTGCCCTTCACAAAAATTTGGTGT 1843  
Qy 1878 TCGCAGTATTTCTAGACGATCTTTTACGACACCCACCGGATCTCTGGAGTATCTTAAATTTGT 1937  
Db 1844 TACAAGTCTCTTTTGATGATATGGCTGACATCTTTGCAACACATAGATGAATTTGAAAGTT 1903

Qy 1938 TCTCTGAAGCGGTCGAGATGGGATATCTCTGTCTGGATAGCGTTCTGGGATATATCAGT 1997  
Db 1904 TCATCTGAGGAGTAAGAGATGGGATACATCTTTGCTACATGAGATTCAGAGTGT---A 1960  
Qy 1998 TGAAGATTTGCTTCTAGGCTGTACAAACACAGTGAATGGATTTGGAAAGATGACTCA 2057  
Db 1961 TGAACATCTGCTTTAAAGTTTGGTTCAAATTAATGAAGAAAGTAAATTAATGATGGTTA 2020  
Qy 2058 AGGACAGGCGCTGATGCTGTGGCTATCTTCGAAAGATATGGAGGCTTGTCTCGCAT 2117  
Db 2021 AGGTACAAGGAGCTGACATGCTCGCTCACATAAGAAACCTCTGGGAGTTGTACTCAAT 2080  
Qy 2118 CGTATACCAAGAACCCGGAATGGTTCGGCAGCAAAAGTATGTGCCGACATTTCAACGAATATG 2177  
Db 2081 GTTATGTACAAGAAAGGAGTGGCTTGAAGCGGGTATATACCAACTTTTGAAGAGTACT 2140  
Qy 2178 TGGAAATGCCAAAGTGTCCATGACATTCGACAGTGTCTTAACACTCAATCTTTTCA 2237  
Db 2141 TAAAGACTTATGCTATATCAGTAGGCTTGGACCGGTGACCCCTACAAACCAATACTACTAA 2200  
Qy 2238 CTGGAGAAATTAATCTCTGATTACATTTTACAGCAAGTAGACCTTCGGTCCAAATTTCTGC 2297  
Db 2201 TGGGTGAGCTTGTGAAGATGATGTTGTTGAGAAAGTGCACCTATCCCTCAATATGTTTG 2260  
Qy 2298 ATCTTGTGCTTTGACTGGAAGCTTAATCAATGACACCAAGACTTACAGGCGGAGAGAA 2357  
Db 2261 AGCTTGTATCTTGAAGTGGGACTTAACAAACGACACCAAAACATATCAGGCTGAAAGG 2320  
Qy 2358 ACCGTGGTGAATTTGTTTCCAGCGTACAGTGTACATGAGGGAATAATCCGAGTGCACAG 2417  
Db 2321 CTCGAGGACCAACAGCCCTCAGGCATAGCATGATGATTAAGAGTATATCAGGAGCACTG 2380  
Qy 2418 AGGAAGAAAGCTCTAAGTCAATGTTTATGTTATCATCGAACACGACTGAAGGAATTTGAAAT 2477  
Db 2381 AGGAAGATGCCATTAAGCACATATGTCGTTGTTGATCGGGCTTGAAGAAGCAAGCT 2440  
Qy 2478 GGGAGTTGGCCAAACCCAGGAGCAATGCCCAATTTGTTGTGTGAGAGACTGCTGTTCACA 2537  
Db 2441 TTGAATATTTCAACACCATCCAAATGATATCCCAATGGGTTGCAAGTCTTTATTTTAACC 2500  
Qy 2538 CTGCAAGAGTGTGAGCTGTTTATATGTACAGAGATGGCTTTGGTATCTCTGACAAAG 2597  
Db 2501 TTAGATTGTTGTCCAATCTTTTACAGTTTATAGATGGTACGGAATCGCCAATGAGG 2560  
Qy 2598 AGATGAAGACCATGTACGCCGAATCTTTTTCGATCC 2634  
Db 2561 AGATTAAAGGACTATATAGAAGAAAGTTTATATTGATCC 2597

## RESULT 11

AAF73412  
ID AAF73412 standard; cDNA; 2186 BP.

XX AAF73412;

XX 30-APR-2001 (first entry)

DE Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 66.

XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.

XX Abies grandis.

XX WO200107565-A2.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-US020264.

XX 26-JUL-1999; 99US-00360545.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX PA Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX PI WPI; 2001-182782/18.  
 XX DR P-PSDB; AAB69391.  
 XX PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX PS Claim 13; Page 151-154; 175pp; English.  
 XX CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX SQ Sequence 2186 BP; 722 A; 440 C; 472 G; 552 T; 0 U; 0 Other;  
 Query Match 16.7%; Score 450.4; DB 4; Length 2186;  
 Best Local Similarity 57.6%; Pred. No. 4.8e-111;  
 Matches 899; Conservative 0; Mismatches 631; Indels 30; Gaps 4;  
 QY 1081 TGATCTATTGTAACGCTCTGGGCTGCGATACAGTTGAACGCTTGGGATCGCTA 1140  
 DB 390 TGAATCTATTGTAACGCTTGGGATGCGATAGCGTTGGAACGTTGGGGATTGATAGACA 449  
 QY 1141 TTTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGTACTTGGGACGCCGAAAG 1200  
 DB 450 TTTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGTACTTGGGATGATAGACA 506  
 QY 1201 AGCGTGGGATGGGCAAGATGCAATCTCTGATGATGCGATGACACATCGCATGGGTCT 1260  
 DB 507 AGGTATTGATGGGATGAGATAGTGTCTTCTGATGTCAACTGCACTGCTCGGGGTT 566  
 QY 1261 TAGAATCTTGAGATTCATCGATACATGATCTTCAAGATGTTCTGGAGATTTTCAGAGA 1320  
 DB 567 TCGAATCTTTCGCTACACGATACAGTGTCTCTCAGAGTTTGAAGATTTCAAGA 626  
 QY 1321 CGAAGAGGAGACTTCTTTTGTCTTTCGCGTCAAGCGAAATGCTGTGACCGATAATCT 1380  
 DB 627 CCAAAATGGGCACTTTCATCTCTCTAGTACAAAGAGAGACATCAGAACCGTTCT 686  
 QY 1381 TAACCTTTATAGATGTTCAAGATATGTTTTCGGGAGAAAGATTAATGGAAGAGCTAA 1440  
 DB 687 GAATTTATATCGGGCTTCTTTCTATGCTTCTTCTGGGAGAAAGTTATGGAAGAGGCTGA 746  
 QY 1441 GACCTTTCATCAAAATCATCTCCAAATGCTCTTGGCCAAAGCAACGCAATTTGATGAAGTG 1500  
 DB 747 AATTTTCTCTTCAAGATATTGAAAGAGCGGTCGCAAGATTTCCGGTCT- 796  
 QY 1501 GGCTGTCAAGAGGATCTTCTCGAGAGGTGGAGTATGCTTATAAGATATCCGTGGCATAG 1560  
 DB 797 -----CCAGTCTTTCACAGAAATAGACTACACTTTTGAATATGTTGGCACAC 845  
 QY 1561 AAGTATGCCAAGATTGGAGCAAGATTTACATAGAGCAATTTGGATCAAAATGATGCTG 1620  
 DB 846 AATATATGCCAAGATTGGAAACAAGGAATTAATCTTAGATGATTTGG-ACATCTACCAG 902  
 QY 1621 GCTGGGGAGACTGTGTATAGATGCTATATGTCAGCAACGAAATATTTGGAGTGGC 1680  
 DB 903 TCCATGGCTCAAGAGAAAGAGCGCAATATCTGGACAGCAAGAGCTTTTAGAATCGC 962  
 QY 1681 CAAATTTGACTTCAATATGTCAGGCTTTACACCAAGAGGAGACTCAACACATTTGTCAG 1740  
 DB 963 AAAATTTGGAGTTCAACATCTTCACTCCCTTCAACAGAGGAGTTACAGTATCTCTCCAG 1022

QY 1741 CTGGTGGAGAGATCGGATTTCAATGATCTTACATTTCACCGCCAGCGGCTCTGGAAAT 1800  
 DB 1023 ATGGTGGATACATTTGGGTTTGGCTGAACCTGTTTGGTGGCATCGTCACTGGGAATA 1082  
 QY 1801 GTATTTCTCAGTGGCGGTTAGTATGTTTGAAGCAGAAATCGCTGCTGTGTAAGATTGCTTA 1860  
 DB 1083 CTACACCTTGAGCTCTTGGCATTTGCGACTGAGCCCAACATTTCTGCAATCAGATTGGGCTT 1142  
 QY 1861 TGCAGAGACTTCTTGGCTCGCATTTATCTAGACGATCTTTAGCAGACCCACCGATCTCT 1920  
 DB 1143 TGCAGAAACGCTGTCTATCATCGGTTCTGGACGATATCTACGACACTTTTCGGAACGAT 1202  
 QY 1921 GGTATCTTAATATTTCTTCTGAAGCGGTGCGAAGATGGATATCTCTGTCTGGTAG 1980  
 DB 1203 GGTATGAATTCGAACCTCTTCAAGAGGAGCTTAGGATGGAATCCGTCGGAGAAAGAACG 1262  
 QY 1981 CGTTCCGGATTAATCAGCTTCAAAAGTTGCTTCTTAGGGCTGTACAAACACAGTGAATGAT 2040  
 DB 1263 CTTCCAGATATA---TGAAGAAATCTACATGGCACTTACGAGGCTTTACTGACAT 1319  
 QY 2041 TGGAAAGATGGAATCAAGGAAACAGGCGGTGATGTCTGGGCTATCTTTCGAAAAGPATG 2100  
 DB 1320 GCGCGAGAGGCGAGAGAGACACAGGCGGAGACAGCTCAATATATGCTAGAAAGGCTTG 1379  
 QY 2101 GNGGCTTGTCTCGCATCTATACAAAGAGCGGAATGGTGGCAGCAAGTATGTGCC 2160  
 DB 1380 GGAAGTTTATCTTGAATTCGTATACAGAAAGCAAAAGTGGATCGCCAGCGGTTATCTGCC 1439  
 QY 2161 GACATTCACAGCAATATGTGCAAAATGCCAAAGTGTCCATAGCACTTTCGACAGCTCGTACT 2220  
 DB 1440 AACTTTCGAGGAGTACTTAGAGAACGGAAGTTAGCTCTGGTCTATCGTCAGCGGCAAT 1499  
 QY 2221 AAATCAATCTTTTTCATCTGGAGAAATTAATCTCTGATTAATTTTACAGCAAGTAGACCT 2280  
 DB 1500 GACACCCCTCTGACATTTGACGCTTCTGATGAGCTTCTGAAGGGAATAGATTT 1559  
 QY 2281 TCGGTCAAAATTTCTGCATCTTGTGTTGTTGACTGGACGACTAATCAATGACACCAAGAC 2340  
 DB 1560 TCCATCGAGATTTAATGATTTGGCATCTTCTCTTTAGACTAAGAGGTGACACAGATG 1619  
 QY 2341 TTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGGTACAGTGTACATAGGGA 2400  
 DB 1620 CTACAGGCGAGACAGGACCGAGGAGAGAGGCTCAGCATATCGTGTTCATGAAAGA 1679  
 QY 2401 AAATCCGAGTGCACAGAGAGAAAGCTCTAAGTCTATGTTTATGTTATCATTCGACAAACGC 2460  
 DB 1680 CAATCCCGGATTAACAGAGAGAGATGCTCTCAATCATATCAATGCCATGATCAACGACAT 1739  
 QY 2461 ACTGAAGGAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAG 2520  
 DB 1740 AATCAAGAAATTTAAATTTGGGAACCTTCTCAACCCGATAGCAATATTTCCATGCTGCACG 1799  
 QY 2521 AAGACTGCTGTTCAACACTGCAAGAGTGTATGCACTGTTTATATGTACAGAGATGGCTT 2580  
 DB 1800 GAAACATGCTTATGAGTAACCCAGACTTTCACCAACTTTACAAATATAGAGATGGCTT 1959  
 QY 2581 TGGTATCTCTGACAAAGAGATGAAAGACCATGTGACCGCAACTCTTTTCGATCTGTGGC 2640  
 DB 1860 CAGCGTTGCCACTCAAGAAACGAAAGTTTGGTGGAGAGAACGGTCTCTTGAACCGATGCC 1919

RESULT 12  
 AAF73421  
 ID AAF73421 standard; cDNA; 1890 BP.  
 XX  
 AC AAF73421;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 DE Grand fir monoterpene synthase coding sequence SEQ ID NO: 77.  
 XX  
 KW Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;



terpinolene synthase; insect resistance; nutrition; BB.  
 XX Abies grandis.  
 OS WO200107565-A2.  
 XX PD 01-FEB-2001.  
 XX PF 24-JUL-2000; 2000WO-US020264.  
 XX PR 26-JUL-1999; 99US-00360545.  
 XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
 XX WPI; 2001-182782/18.  
 DR P-PSDB; AAB69393.  
 XX PT New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.  
 XX PS Claim 18; Page 163-165; 175pp; English.  
 CC The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX SQ Sequence 1890 BP; 550 A; 381 C; 441 G; 518 T; 0 U; 0 Other;  
 Query Match 15.4%; Score 416; DB 4; Length 1890;  
 Best Local Similarity 57.3%; Pred. No. 9.3e-102;  
 Matches 899; Conservative 0; Mismatches 620; Indels 51; Gaps 6;  
 1081 TGATCTATTGTTGACCGCTCTGGGCTGTCGATACAGTTGACGCTTGGGAATCGATCGCTA 1140  
 366 TGATCTCATTCACCGCTTTAATGTCGATACAGTTGACGCTTGGGAATCGATCGCTA 425  
 1141 TTTCAGAAAGAAATCAAAGAATCTCTGGATACGTTTATAGTACTGGGACGCCGAAAG 1200  
 426 TTTCAGAAATGAGATAAAGCAGCACTAGACTATGTTTACAGTTATTGGAA---CGAAAA 482  
 1201 AGCGTGGGATGGGCAAGATGCMATCCTATTCTGATGTCGATGACACTGCCATGGGTCT 1260  
 483 AGGCATTGGCAGTGGAAAGTGATAGTGGTGTGCTGATCTCAACTCACTGCCCTGGGGTT 542  
 1261 TAGAATCCTGAGACTTTCATGGATACAAATGATCTTCAGATGTTCTCGAGAAATTCAGAG- 1319  
 543 TCGAATCTTCGACTACAGATACAGTGTCTTCAGATGTTGTTGGAACACTTCAAAGA 602  
 1320 -----ACGAGAAAGGAGACTCTTTTGTCTTGGCGGTCAAACGCAAAATGGTGTGACCGA 1374  
 603 GGAGAGAGGAGAGGGGAGTTTGTATGTTTGGGCCATCCAAACAGAGAAAGAGATAAAAAAG 662  
 1375 TAATCTTAACCTTTATAGATGTTTCAAGATGATGTTTTCGGGAGAGAAAGATATGAAGA 1434  
 663 CGTTCTGMAATTTAATTTTCGGGCCCTCCCTCATTTGCTTTCCTGGGAGAGAGTTATGAAGA 722  
 1435 AGCTAAGACCTTCACCTACAAATCATCTCCAAAATGCTCTTTGCCAAAACCAACGCAATTTGA 1494  
 723 GGCTGGAATCTTCTTAATAATATTT-----AAAGAACCCCTTACA 764  
 1495 TAAGTGGGCTGTCAAAGAGATCTTCTCGAGAGAGTGGAGTATGCTATTAAGTATCCGTG 1554  
 765 AAATATTGCTGTCTCCA---GTCTTTACAGAGAGATAGATGCTTCTCGAGAGATGGTTG 821  
 1555 GCATAGAGTATGCCAAGATTGGAGGCAAGAGTTACATAGAGCAATTTGGATCAAATGA 1614

822 GCAAAACAATATGCCAAGATTGGAAAACAAGGAACATACATGATGTATTGGGA----- 873  
 1615 TGTCTGGCTGGGGAAGACTGTGTATAAGATCTCTATATGTAGCAACAGAAAAATATTTGGA 1674  
 874 -----GAGAACGATCGTATGAGACGTTATATATGACATGGAGAACTTTTAGA 923  
 1675 CTGGCCAAATTTGGAATTTCAATATGGTCAGCGCTTACACAAAGAGAGACTCAACACAT 1734  
 924 AATTGCAAAATTTGGAGTTCAATATTTTCTCTCTCAACAGAGAGAGACTAAAAAGACT 983  
 1735 TGTGAGCTGGTGGAGAGATCGGAGTTCAATGATCTTACATTCACCCGCGAGCGGCTGT 1794  
 984 CTCAGATGGTGGAAAGATTCCGGTTCTCTCACCTGACATTTTCTCGGATCGTCATGT 1043  
 1795 GGAATGTATTTCTCAGTGGCGGTAGTATGTTTGGCCAGAAATTCGCTGTGTAGAAAT 1854  
 1044 GGAATCTACGCTCTGGCATCTTGCAATTGAACTGATCGCAACATTCGGGATCAGACT 1103  
 1855 TGCTATGCCAAGACTCTTCTGCTCGCAGATTATCTAGACGATCTTTAGGACACCCACGG 1914  
 1104 CGGCTTTGCCAAAATGTGTCTATCATCGGTTTGGACGATATATACGACACTTTGG 1163  
 1915 ATCTCTGATGATCTTAATTTGTTCTGAAAGCGTCCGAAAGATGGGATATCTCTGTGCT 1974  
 1164 AACATGGAGGAGCTGGAACTCTTCTCTGACGATTTAAGAGATGGGATCCGCTCTGCCAC 1223  
 1975 GGATAGCGCTTGGGATAATCAGTTGAAAGTTTGGCTTCTTAGGGCTGTACAAACAGTGAA 2034  
 1224 AGATTGCTTCCAGAGTATA---TGAAAGGTTGTACATGGTGGTTTACGAAACCGTAAA 1280  
 2035 TGGATTGGAAAGATGAGACTCAAGGAAACAGGCGGTGATGTCTGGGCTATCTTCGAAA 2094  
 1281 TGAATTTGCTCGAGAGGCGAGCAAGTCTCAAGGCGGAGAGACGCTCAAGATGCTCGACG 1340  
 2095 AGTATGGGAGGCTTGTCTCGCATCGTATACCAAGAACCGGAATGCTCGGACGACAAAGTA 2154  
 1341 AGCTTGGAGGCTATCTTGTATGATGATGAAAGAGCTGAGTGGATCTCCAGTGGTTA 1400  
 2155 TGTGCGCACTTCAACGAATATGTGGAAATGCCAAAGATGTCATAGCACTTGGCAGACT 2214  
 1401 TCTGCGCAAGTTGAGGAGTACATGGAGACCAAGCAAGTTAGTTTGGTTATGCGCATAT 1460  
 2215 CGTACTCAACTCAATCTTTTCTCAGTGGGAATTTACTTCTGATTTACATTTTACAGCAAGT 2274  
 1461 CGCATTTGCAACCCATCTCTCATATGATGTTCCCTTACTCACCACATCTCTCGAGAAAT 1520  
 2275 AGACCTTCGGTCCAAATTTCTGCATCTTGTCTTTGACTGGACGACTAATCAATGACAC 2334  
 1521 AGACTTTCATTTGAGTTTAAATGACTTAATATGTTTCCATCTCTCGACTTAAAAATGACAC 1580  
 2335 CAAGACTTACAGGCGGAGAGAAACGTTGGTGAATTTGTTTCCAGCGTACAGTGTACAT 2394  
 1581 TCGCTGCTCAAGAGGCGGACAGGGCCGTTGGAGAAAGCTTCTGTTATATCGTGTATAT 1640  
 2395 GAGGGAATATCCGAGTGCACAGAGAGAAAGCTCTAAGTCTATGTTTATGTTATCATCGA 2454  
 1641 GAAAGAAATCTTGGATCAACAGAGAGAAAGTCTATCATCATATCAACGCTATGTCAT 1700  
 2455 CAACGCACTGAAAGAAATGAAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCCATGTG 2514  
 1701 TAACTTAAATCAAAAGAGTGAATTTGGAGCTTCTCCGACAGGAGCGGACCGCTCATATTGC 1760  
 2515 TGTGAGAAAGCTGCTGTTCAACACTGCAAGAGTGTAGTGTGTTTATATATGATACAGAGA 2574  
 1761 TTGCAAGAAACAGCTTTTGGACATCTCAAGGTTCCCTTTCAGGCTACAAATACCGAGA 1820  
 2575 TGGCTTTGTTGATCTCTGCAAAAGAGATGAAAGCAATGTGACGCCAACTCTTTTCGATCC 2634  
 1821 TGGTTCCAGGTTGCCCAACAGGAAACCAAGAAATTTGGGTGAGGAGAAAGTCTCTGAGTC 1880  
 2635 TGTGGCGGTAG 2644

Db	1881	TGTGCTTTG	1890	Qy	1261	TAGAACTCTGAGACTTTCATGGATACAATGTATCTTTCAGATGTTCTGGAGAAATTCAGAGA	1320
				Db	596	TCGAACCTCTTCGATTATACACGGGTACACTGTATCTCCAGAGGTTTAAAGCTTTTTCAGA	655
RESULT 13				Qy	1321	CGAGAAAGGAGACTTCTTTTTCCTGCTTTCGCGTCAAAAGCAAAATGGTGTGACCGGATAATCT	1380
AAAX08663				Db	656	TCAAAATGGACAGTTTGTATGCTCCCGGTCAGACAGAGGGTGAGATCAGAAGCGTTCT	715
AC	AAAX08663;			Qy	1381	TAACTTTTATAGATGTTTCAAGTATGTTTTCGGGAGAAAAGATATATGGAAGAGCTAA	1440
XX	27-SEP-1999 (first entry)			Db	716	TAACTTTATATCGGGCTTCCCTCATTTGCTTCCCTGTTGAGAAAGTTATGGAAGAGCTGA	775
XX	Grand Fir monoterpene synthase clone AG3.48.			Qy	1441	GACCTTCACTAGAAATCATCTCCAAATGCTCTTGGCCAAAACCAACGCCATTTGATAGTG	1500
XX	Myrcene synthase; limonene synthase; pinene synthase; flavour;			Db	776	AATCTTTCTCCACAGATATTTGAAAGAGCTCTACAAAAGATTTCCAGTCTC-----	826
KW	monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;			Qy	1501	GGCTGTCAAGAGGATCTTCTCGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAG	1560
KW	probe; ss.			Db	827	-----CGCTCTTTCAAGAGATAAAGTTTGTATGGAATATGCTGGCACAC	874
OS	Abies grandis.			Qy	1561	AAGTATGCCAAGATTGGAGGCAAGAAGTTACATAGAGCAATTTG---GATCAAAATGATG	1617
FH	Key	Location/Qualifiers		Db	875	AAATTTGCCAAGATTGGAAGCAAGAAATTTACATAGACACACTTTGAGAAAGACACACG	934
FT	CDS	57..1940		Qy	1618	CTGGCTGGGGAAGACTGTGTATAGATGCTATATGTAGCAACGAAAAATATTTGGAGCT	1677
FT	/*tag= a	/product= "Monoterpene synthase"		Db	935	ATGGCTCAATAAATAATGCTGGGAAGAGCT-----TTTAGAAT	973
XX	WO9902030-A1.			Qy	1678	GGCCAAATTTGGACTTCAATATGCTGCAGGCTTTACACCAAAAGAGACTCAACACATTTGT	1737
XX	21-JAN-1999.			Db	974	TGCAAAATTTGGAGTTCAATATATTTAACTCTTACAAACAAAGGAATTTACAAATCTTTT	1033
XX	10-JUL-1998; 98WO-US014528.			Qy	1738	CAGCTGTGGAGAGAAATCGGGATTCAATGATCTTACATTCACCCGCGAGGGCTGTGGA	1797
XX	11-JUL-1997; 97US-0052249P.			Db	1034	GAGATGTGGAAAGAGTGGGATTTGCTTAATTTGACATTTGCTCGGCATCGTCATGTGA	1093
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.			Qy	1798	AATGTATTTCTCAGTGGCGGTTAGTATGTTTGGAGCCGAAATTCGCTGCTGTGTAGAAATGC	1857
XX	Bohlmann J, Steele CL, Croteau RB;			Db	1094	ATTCTACACTTTGGGCTCTTGTATTTGCCATTGACCCCAAAACATTTCTGCATTTCAAGCTAGG	1153
XX	WPI; 1999-120396/10.			Qy	1858	CTATGCCAAGACTTCTTGCCCTCCGAGTTATTTCTAGACGATCTTTTACGACCCACCGCATC	1917
XX	P-PSDB; AAW85710.			Db	1154	CTTCGCCAAATGTGTCTATCTTGTACAGTTTGGAGCATTTTACGACATTTTGGAAAC	1213
XX	New isolated gymnosperm monoterpene synthase DNA - obtained from Grand			Qy	1918	TCGGATGATCTTAAATTTGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGTGGA	1977
XX	fir (Abies grandis), used to provide plants with modified production of			Db	1214	GATTCAGAGCTTGAACCTCTTCACTCTGCAATTAAGAGATGGAATTCATCAGAGATAGA	1273
XX	monoterpenes, e.g. myrcene, limonene or pinene.			Qy	1978	TAGCGTTTGGGATAAATCAGTTGAAAAGTTTGTCTTCTTAGGCTGTACAAACAGTGAATGG	2037
XX	Example 3; Page 104-107; 121pp; English.			Db	1274	ACACCTTCCAGAAATATA---TGAATGTGTGTACATGCTGTGTGTTGAAACTGTAAATGA	1330
XX	Nucleotide sequences encoding myrcene synthase, limonene synthase and			Qy	2038	ATTGGAAGAAGATGGACTCAAGGCAAGGCCGTGATGTCTGGGCTATCTTCGAAAGT	2097
XX	pinene synthase from Grand fir may be incorporated into any organism			Db	1331	ACTGACACGAGAGCGGAGAGAACTCAAGGGAGAAACACTCTCAACTATGTTCGAAAGGC	1390
XX	(e.g. intact plant, animal, microbe), or derived cell culture that			Qy	2098	ATGGAGGGCTTGTCTCGCATCGTATACCAAGAGCGGAATGGTGGCGAGCAAGATATGT	2157
XX	produces geranyl diphosphate for the production of the aforementioned			Db	1391	TTGGAGGCTTATTTTGTATTCATATATGGAAGAACAAATGGATCTCTAATGTTATCT	1450
XX	enzymes or their products. The sequences when expressed in transfected			Qy	2158	GCCGACATTCACGAATATGTGAAAATGCAAAAGTGTCCATAGCACTTGGCAGTCGT	2217
XX	cells may also be used for the production or modification of flavour and			Db	1451	GCCAAAGTGTGAAGAGTACCATGAGAATGGGAAAGTGAGCTCTGCATATCGCGTAGCAAC	1510
XX	aroma properties, improvement of defense capability, and the alteration			Qy	2218	ACTAAACTCAATCTTTTCTACTGGAGAAATTAATCTTCTGATTAATTTTACGCAAGTAGA	2277
XX	of other ecological interactions mediated by myrcene, limonene, pinene,			Db	1511	ATTGCAACCCATCTCACTTTGAAATGCAATGGCTTCTCTGATTAATCTTTGAAGGAAATGA	1570
XX	or their derivatives. In particular they can be used for the production			Qy	2278	CTTTCGCTCCAAATTTCTGCATCTTGTGTCTTTGACTGGACGACTAATCAATGACACAA	2337
XX	of plant seeds for the extraction of oil or meal			Db	1571	TTTTCCATCCAGGTTCAATGATTTGGCATGCTCTCTTCCTTCGGCTAGAGGTGACACAG	1630
SQ	Sequence 2205 BP; 703 A; 412 C; 463 G; 627 T; 0 U; 0 Other;						
Query Match	15.28;	Score 412.2;	DB 2;	Length 2205;			
Best Local Similarity	56.8;	Pred. No. 1.1e-100;					
Matches 887;	Conservative 0;	Mismatches 623;	Indels 51;	Gaps 5;			
Qy	1081	TGATCTATTTGAACGCTCTGGGCTGTCGATACAGTTTGAACGCTTGGGAATCGATCGCTA	1140				
Db	419	TGATCTCATGCAACGCTTTGGATAGTCGATAGCGTTGACGTTTGGGATAGTAGACA	478				
Qy	1141	TTTCAAGAAAGAAATCAAGAAATCTCTGGATTAGCTTTATAGTACTGGGAGCCGGAAG	1200				
Db	479	TTTCAAGAAACGAGATAACATCAGCTCTGGATTATGTTTTCGTTACTTGGGA---GGAAA	535				
Qy	1201	AGGCTGGATGGCAAGATGCATCTTATCTCGATGTCGATGACACTGCCATGGGTCT	1260				
Db	536	CGGCATTTGGATGTGGGAGAGACAGTATTTGTACTGATCTCAACTCAACTCGTTGGGTT	595				



Db 1391 TTGGAGGCTTATTTTGAATCATATATGGAAGAGCAAAATGGATCTCTAATGGTTAICT 1450  
 Qy 2158 GCCGACATTCACGAATATGTGGAATATGCCAAAGTGTCCATAGCACTTGGCAGTCGT 2217  
 Db 1451 GCCAACGTTTGAAGAGTACCATGAGATGGAAAGTGCAGCTCTGCATATCGGTAGCAAC 1510  
 Qy 2218 ACTAAATCATCTTTTTCACGTGAGAAATTAATCTTCTGATTAATTTTACAGCAAGTAGA 2277  
 Db 1511 ATTGCAACCCATCCTCACTTTGAATGATGATGGCTTCTCTGATTAATTTGAAGGGAATTGA 1570  
 Qy 2278 CTTTCGGTCCAAATTTCTGCATCTTGTGCTTTTGAATGATGATGGCTTCTCTGATTAATTTGAAGGGAATTGA 2337  
 Db 1571 TTTTCCATCCAGTTCAATGATTTGGCATCTCTCTTCTTGGCTACGAGTGACACG 1630  
 Qy 2338 GACTTACAGCCGAGAGAAACCGTGTGTAATTTTCCAGCGGTACAGTGTCTACATGAG 2397  
 Db 1631 CTGCTACAGGCCGATAGGATCGTGTGGAAGAGTCTCGTATATCATGTTATATGAA 1690  
 Qy 2398 GGAATATCCGAGTGCACAGAGGAGAAAGCTCTAAGTCATGTTTATGTTATCATCGACAA 2457  
 Db 1691 AGCAATCTCTGGATCAACCGAAGAGATGCCCTCAATCATATCAATGCTTCAATGA 1750  
 Qy 2458 CGCACTGAAGGAATTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGT 2517  
 Db 1751 CATATCAAGAAATTAATTTGGAACTTCTAAGATCAACAGCAATATTCATGCTGGC 1810  
 Qy 2518 GAGAAGACTGCTGTTCAACACTGCAAGAGTGAATGATGATGTTTATATGATACAGATGG 2577  
 Db 1811 CAAGAAACATGCTTTTGACATAACAGAGTCTCCACCATCTCTACATATATGAGATGG 1870  
 Qy 2578 CTTTGGTATCTTGACAAAGAGTGAAGACCATGTGAGCGGAATCTCTTTTCGATCCTGT 2637  
 Db 1871 CTTTAGTGTGTCACCAAGGAAACAAAAAATTTGGTTATGGAACACTCTCTTGAATCTAT 1930  
 Qy 2638 G 2638  
 Db 1931 G 1931

RESULT 15  
 ID AAA38933 standard; DNA; 1865 BP.  
 AC AAA38933;  
 XX 25-AUG-2000 (first entry)  
 DT Grand fir delta-selinene synthase DNA sequence SEQ ID NO:47.  
 XX  
 DE  
 XX  
 KW Synthase; protein co-ordinate data; active site; modification; terpenoid;  
 KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
 KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
 KW defensive agent; pigment; antitumour; steroid hormone;  
 KW signal transduction pathway; bile acid; affinity purification;  
 KW photoreceptor; enzymatic synthesis; nutrient supplement;  
 KW immunological reagent; ds.  
 XX  
 OS Abies grandis.  
 XX  
 XX WO200017327-A2.  
 PN  
 XX  
 XX 30-MAR-2000.  
 PD  
 XX  
 XX 17-SEP-1999; 99WO-US021419.  
 PP  
 XX  
 XX 18-SEP-1998; 98US-0100993P.  
 PR 22-APR-1999; 99US-0130628P.  
 PR 23-AUG-1999; 99US-0150262P.  
 XX  
 XX (KENT ) UNIV KENTUCKY RES DEPT.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 XX

PI Chappell J, Manna KR, Noel JP, Starks CM;  
 XX WPI; 2000-292839/25.  
 DR P-PSDB; AAY90854.  
 XX  
 PT Novel terpene synthase enzymes, useful for producing terpene  
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
 PT enzymes by specific amino acid alterations.  
 XX  
 PS Disclosure; Page 420-423; 450pp; English.  
 XX  
 CC The present invention describes an isolated terpene synthase (I)  
 CC comprising a region with at least 20% identity to region 265-535 of a 548  
 CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
 CC -carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom ) within tabulated ranges, and have an ordered  
 CC arrangement of R groups (defining as side chains), excluding specific  
 CC tabulated arrangements (tables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 1865 BP; 580 A; 340 C; 434 G; 511 T; 0 U; 0 Other;  
 Query Match 15.2%; Score 411.2; DB 3; Length 1865;  
 Best Local Similarity 57.4%; Pred. No. 1.0e-100; Indels 48; Gaps 7;  
 Matches 903; Conservative 0; Mismatches 623;  
 Qy 1079 CTTGATCTATTGAAACGCTCTCGGCTGTGATACAGTCTTGAACGCTTGGGAATCGATGC 1138  
 Db 190 CATGATTTAATCAACGCTCTTCAGATCGTTGACACTTTTGAATGCTCGTGAATCGATAGA 249  
 Qy 1139 TATTTCAGAAAGAAATCAAGAAATCTC---TGGATTACGTTTATAGTACTGGGACGCC 1195  
 Db 250 CATTTTGAACACGAAATACAAACAGCTGCTTTAGATTACGTTTACAGATGCTGGAACGAA 309  
 Qy 1196 GAAAGAGCGCTGGATGGGCAAGATGCAATCTTCTGATGTCGATGACACTGCCATG 1255  
 Db 310 AAAGGTATCGGGAGGGATCAAGAGATCTCTCAGCAAGATCTGAACGCTACGCTTTA 369  
 Qy 1256 GGTCTTAGAATCTCGAGACTTCATGGATACAATGTATCTTTCAGATGTTCTGGAGAATTC 1315  
 Db 370 GGATTTTCGCGCTCTCGAGCTGCATCGATATAACGATATCGTCAAGTGTGTTGAAGAATTC 429  
 Qy 1316 AGAGACGAGAAAGAGACTCTTTTGTG---CTTTGCCGCTCAACGCAAAATTTG----- 1365  
 Db 430 AAGGATGAAACCGGAAGTTCTTCTGCAACTTTTCTGCAAGTCTTGAAGAGGAGAGATAAA 489  
 Qy 1366 --TGTGACCGATAATCTTAACCTTTTATAGATGTTTCAAGTATGTTTTCGGGAGAGAAAG 1423  
 Db 490 CAAGTGAGAAAGCATGTTGTCTGTTACTTTCGAGCTTCAGAGATTCGTTTCCGGAGAAAAA 549  
 Qy 1424 ATAAATGAAGAAGCTTAAGACTTTCACATCAAAATCATCTCCAAATGTCTTTCGCCAAAAAC 1483  
 Db 550 GTGATGAAGAGGCCAAGGCATTCACAAAGAGATATCTAAACCAAGTTTTCAGCTGGACAC 609  
 Qy 1484 AAGCATTTGATAGTGGGCTGTCAAGAGGATCTTCTCGAGAGAGGTGGAGTATGCTATA 1543  
 Db 610 GGG-----GATGTGACTGACGTGGATCAAAAGCCTTTTGGAGAGAGGTGAAGTACGCAATTG 663  
 Qy 1544 AAGTATCCGTGGCATAGAACTATGCCAAGATTTGGAGGCAAGAGTTACATAGACCAATTT 1603

Search completed: August 24, 2004, 21:15:21  
Job time : 1016.04 secs

Db	664	GAGTTTCCATGGCATTCGATGTCGCCGAGATGGGCAAGGAGCTTCTCGAATATAT	723
Qy	1604	GGATCAAAATGATGTCGTGGGGAAGACTGTGTATAAGATGCTATATGTGAGCAACGAA	1663
Db	724	GGACACAAACCATTCGTGGCTCAAGTCGAATAT-----CAACCAA	762
Qy	1664	AAATATTTGGAGCTGGCCAAATTCGACTTCAATATGTTGCAGGCCCTTACACCAAAAGGAG	1723
Db	763	AAAATGTTGAAGTTAGCCAAATTCGACTTCAATATTCGCAATGCAAAATCACAAGGAG	822
Qy	1724	ACTCAACACATTCAGCTGGTGGAGAGAATCGGGAATCAATGATCTTACATTCACCCGC	1783
Db	823	ATACAGTTTATTTACAGGTGGTGGAGAGACTCGGGTATATCGCAGCTGAATTTCTATCGA	882
Qy	1784	CAGCGGCTGTGGAAATGATTTCTCAGTGGGGGTAGTATGTTTGGAGCCAGAAATTCGCT	1843
Db	883	AAGCGACACGTGGAATATTAATCTTGGGTGTTATGTCATTTTGGCCGAGTTCTCT	942
Qy	1844	GCTTGTAGAAATGCGCTATGCCAAGACTTCTTCCCTCGCAGTTTATTTCTAGACGATCTTTAC	1903
Db	943	GAAGTAGAAATTCGCTTCGCCAAACTGCTATCCTGTGTACTGTCTTAGATGACCTCTAT	1002
Qy	1904	GACACCCACGGATCTCTGGATGATCTTAAATTTCTCTGAAGCGGTCCGAAGATGGGAT	1963
Db	1003	GATACGCACGCAACATTCGATGAATCAAAATCATGACAGAGGGAGTGAGACGATGGAT	1062
Qy	1964	ATCTCTGTCTGGTAGCGTTCCGGGATATCAGTTGAAAGTTTGGCTTCTAGGGCTGTAC	2023
Db	1063	CTTTCGTTGACAGATGACCTCCAGAC---TACATTAATAAATTCGATTCAGTCTTCTTC	1119
Qy	2024	AACACAGTGAATGGATTTGGAAAGATGGACTCAAGGAAACAAGGCCGTGATGTCTGGGC	2083
Db	1120	AATACAGTGAATGAATGTAGTTGAAATCGTGAACGGCGGGGATATGACAACC	1179
Qy	2084	TATCTTCGAAAGTATGGAGGGCTTGTCTGGATCGTATACCAAGAGCCGAATGGTCG	2143
Db	1180	ATAGTTAAAGATTCGTGGAAGCGATACATTTGAGTCTTATCTGCAAGACGGGAATGGATA	1239
Qy	2144	GCAGCAAGTATGTGCGGACATTCACAGAAATGTGGAATAATGCCAAAGTCCATAGCA	2203
Db	1240	GCAACTGGACATATTCCTTAAACGATATACATAAGAAACGGCATGGCTAGCTCAGGG	1299
Qy	2204	CTTGGCAGAGTCGTACTAACTCAATCTTTTTCACCTGGAGAAATTAATCTCTGATTCATTT	2263
Db	1300	ATGTGATTTCTAAATTTGAATCCACTTCTTGTGGATAAACTTCTCCCGCACAACTT	1359
Qy	2264	TTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCATCTTGTGCTTTTGACTGGACGACTA	2323
Db	1360	CTGGAGCAAAATACATTTCTCCATCCAGATCCTGGACCTTTAGAAATGACGGSCAGATC	1419
Qy	2324	ATCAATGACACCAAGACTTACCAAGCGCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA	2383
Db	1420	GCCGATGACTTAAAGATTTTCGAGGACGAGAGGAACGGGGGAGATGGCTTCATCTTTA	1479
Qy	2384	CAGTGTACATGAGGGGAAATCCGGAGTGCACAGAGAGAGAGCTCTAAGTCATGTTTTAT	2443
Db	1480	CAGTGTATATGAAGAGAAATCCTGAATCTACAGTGGAAATGCTTTAAATCATAAATA	1539
Qy	2444	GGTATCATCGACCAACGACCTGAAGGAATTCGAATTTGGGAGTTGGCAACCCAGCGAGCAAT	2503
Db	1540	GGCATCTCTTAATCGTTCCTTGGAGAAATTTAATTTGGAGTTTATGAGCAGGATAG---T	1596
Qy	2504	GCCCAATGTGTGAGAGAGCTGCTGTTCAACACTGCAAGAGTATGACAGCTGTTTTTAT	2563
Db	1597	GTCCCAATGTGTTGCAAGAAATTCACCTTCAATATATAGTCGAGGACTTCAATTCATCTAC	1656
Qy	2564	ATGTACAGAGATGGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTGACCGCAACT	2623
Db	1657	AAATACAGAGACGGCTTATACATTTCTGCAAGGAAGTAAGGACCAGATATTCAAATTT	1716
Qy	2624	CTTTTCGATCCTGT	2637
Db	1717	CTAGTCCACCAAGT	1730

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 11:34:34 ; Search time 181.026 Seconds  
(without alignments)  
8292.408 Million cell updates/sec

Title: US-10-041-007-1  
Perfect score: 2705  
Sequence: 1 atattgcctaactgcac.....aaaaaaaaaaaaaaaaaaaaa 2705

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgm2\_6/ptodata/2/ina/5A COMB.seq:  
2: /cgm2\_6/ptodata/2/ina/5B COMB.seq:  
3: /cgm2\_6/ptodata/2/ina/6A COMB.seq:  
4: /cgm2\_6/ptodata/2/ina/6B COMB.seq:  
5: /cgm2\_6/ptodata/2/ina/PCITUS COMB.seq:  
6: /cgm2\_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1091.4	40.3	2861	4	US-09-398-395A-55 Sequence 55, Appl
2	1091.4	40.3	2861	4	US-09-887-586A-55 Sequence 55, Appl
3	1091.4	40.3	2861	4	US-09-895-752-55 Sequence 55, Appl
4	1091.4	40.3	2861	4	US-09-903-012B-55 Sequence 55, Appl
5	1091.4	40.3	2861	4	US-09-900-797-55 Sequence 39, Appl
6	731.6	27.0	2525	3	US-09-234-393-39 Sequence 39, Appl
7	731.6	27.0	2525	4	US-09-865-171-39 Sequence 39, Appl
8	730.2	27.0	2424	3	US-09-234-393-1 Sequence 1, Appl
9	730.2	27.0	2424	4	US-09-360-545-15 Sequence 15, Appl
10	730.2	27.0	2424	4	US-09-865-171-1 Sequence 1, Appl
11	730.2	27.0	2424	4	US-09-398-395A-45 Sequence 45, Appl
12	730.2	27.0	2424	4	US-09-887-586A-45 Sequence 45, Appl
13	730.2	27.0	2424	4	US-09-895-752-45 Sequence 45, Appl
14	730.2	27.0	2424	4	US-09-903-012B-45 Sequence 45, Appl
15	730.2	27.0	2424	4	US-09-900-797-45 Sequence 37, Appl
16	730.2	27.0	2528	3	US-09-234-393-37 Sequence 37, Appl
17	730.2	27.0	2528	3	US-09-234-393-41 Sequence 41, Appl
18	730.2	27.0	2528	4	US-09-865-171-37 Sequence 37, Appl
19	730.2	27.0	2528	4	US-09-865-171-41 Sequence 41, Appl
20	730.2	27.0	2571	3	US-09-234-393-12 Sequence 12, Appl
21	730.2	27.0	2571	4	US-09-865-171-12 Sequence 12, Appl
22	642.2	23.7	2700	3	US-09-315-861-1 Sequence 1, Appl
23	642.2	23.7	2700	4	US-09-398-395A-43 Sequence 43, Appl
24	642.2	23.7	2700	4	US-09-887-586A-43 Sequence 43, Appl
25	642.2	23.7	2700	4	US-09-895-752-43 Sequence 43, Appl
26	642.2	23.7	2700	4	US-09-903-012B-43 Sequence 43, Appl
27	642.2	23.7	2700	4	US-09-593-253-1 Sequence 1, Appl

28	642.2	23.7	2700	4	US-09-900-797-43 Sequence 43, Appl
29	450.4	16.7	2186	4	US-09-360-545-66 Sequence 66, Appl
30	416	15.4	1890	4	US-09-360-545-77 Sequence 77, Appl
31	412.2	15.2	2205	4	US-09-360-545-31 Sequence 31, Appl
32	411.2	15.2	1865	4	US-09-398-395A-47 Sequence 47, Appl
33	411.2	15.2	1865	4	US-09-887-586A-47 Sequence 47, Appl
34	411.2	15.2	1865	4	US-09-895-752-47 Sequence 47, Appl
35	411.2	15.2	1865	4	US-09-903-012B-47 Sequence 47, Appl
36	411.2	15.2	1865	4	US-09-900-797-47 Sequence 47, Appl
37	411.2	15.2	1885	3	US-09-234-393-43 Sequence 43, Appl
38	411.2	15.2	1885	3	US-09-234-393-45 Sequence 45, Appl
39	411.2	15.2	1885	3	US-09-234-393-47 Sequence 47, Appl
40	411.2	15.2	1885	4	US-09-865-171-43 Sequence 43, Appl
41	411.2	15.2	1885	4	US-09-865-171-45 Sequence 45, Appl
42	411.2	15.2	1885	4	US-09-865-171-47 Sequence 47, Appl
43	411.2	15.2	1888	3	US-09-234-393-19 Sequence 19, Appl
44	411.2	15.2	1888	4	US-09-865-171-19 Sequence 19, Appl
45	410.6	15.2	2196	4	US-09-360-545-1 Sequence 1, Appl

## ALIGNMENTS

RESULT 1  
US-09-398-395A-55  
; Sequence 55, Application US/09398395A  
; Patent No. 6468772  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6468772L, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/398,395A  
; CURRENT FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(2606)  
; OTHER INFORMATION: abietadiene synthase  
US-09-398-395A-55

Query Match 40.3%; Score 1091.4; DB 4; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;  
QY 350 GTCCATGAGCAATCAATAAAGAAATTCGAGAGGAAATCCAGTGCATGTTT 409  
DB 324 GTTCAGCATCAGCAGAGACGGTATCGACATTAATATCCGAGATTGAAGATGTTT 383  
QY 410 CAGTCCATGGCGCAGCGGTGAACGAATCCATCTGCATATGATACAGCTTGGTGGCAAGA 469  
DB 384 AGATGTATGGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTTGGGTAGCAAGG 443  
QY 470 ATTCCGTCATTTAGCGCTCTGGTGCACCCCAATTTCCCAAGCCTTCATGATGTTTG 529  
DB 444 ATTCAGCAGTTGATGGCTCTTGACAAACCTCCTCCTCGAGACGGTTGAATGATGTTT 503  
QY 530 AACATCAACTGCCAGATGGCTCGTGGGTGAGGAGTGCAATTTTCTGGCGCTATGACAGA 589  
DB 504 CAAATCAGTTGAAAGATGGGTCTTGGGGTGAGAGTTCTACTTCTTGGCATATGACAGA 563

Thu Aug 26 08:27:49 2004

QY	590	GTGTTTAAACACTCTCGCTGCTCTCTCACTCTCAAAATATGAATAAGGGCGACATCAAA	649	1670	TTGAGCTGCGCCAAATTTGGACTTCAATATGTTGTCAGCGCTTACACAAAAGGAGACTCAA	1729
Db	564	ATAGTGGCTACACTTGGCATGTATTATTACCTTACCTCTGCGCTACTGCGGAGACACAA	623	1641	TTAGAACTAGCGAACTGGAACTGAGTCTCAATAAGTGCAGTCTATACACAAACAGAGCTTCAA	1700
QY	650	GTGAGAAAGGGTGTGATTTGTGAGAAACACATGGAAGAAATGAAGGAGCAAGCTGAC	709	1730	CACATTTGCTAGCTGGTGGAGAGAAATCGGGAATCAATGATCTTACATTCACCCGCGAGCGG	1789
Db	624	GTACAGAAAGGTATTGAATTTCTTCAGGACACAAGCTGGAAAGATGGAAGATGAAGCTGAT	683	1701	GATCTTCGAAGGTGGTGGAAATCATCCGTTTTCACGGATCTGAATTTTCACTCTGAGGCT	1760
QY	710	AATCACAGCCCAAGTGGATTCGAGTCTGTTTCTGCAATGTTAGATGAAGCAAAAAGC	769	1790	CTGTGGAATGTATTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCTGCTGT	1849
Db	684	AGTATAGCCCAAGTGGATTTGAATAGTATTCTGCAATGCTAAAGGAAGCTTAAATC	743	1761	GTGACGGAATATATTTCTCACCGCATCTTATCTTTGAGCCGAGTTTCTTAAGTGC	1820
QY	770	TTGGATTTGGATCTTCTTATACCTCCCTTTCATCTCCCAATCCCAAAAGCGCGAG	829	1850	AGAAATTCCTATGCAAGACTTCTTTCGCTCGAGTATTTAGACGATCTTTAGACACC	1909
Db	744	TTAGGCTTTGGATCTGCTTACGATTTGGCAATCTCTGAAACAAATCATCGAAAGCGGAG	803	1821	AGAGAGTTTATACAAAACCTTCCAATTTTCACTGTTATTTTAGATGATCTTTATGACGCC	1880
QY	830	AAAGAGCTTCAAAAGATTCCTCAATGTTCTTCAATACCACTCAGACGCGTGTCTCTAC	889	1910	CACGATCTCTGGATGATCTTAAATTTGTTCTCAAGCGGTCCGAAGATGGGATATCTCT	1969
Db	804	GCTAAGCTTAAAGAGATTTCCCACTGATGTTCTCTATGCGCTTCCAACAACGTTATTGAT	863	1881	CATGGATCTTTAGACGATCTTTAAGTTTTCACAGATCAGTCAAAAGATGGATCTATCA	1940
QY	890	TCTCTGGAGGTTTGCAGATGTTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGA	949	1970	GTGCTGGATAGCGTTTCGGGATAATCAGTTGAAAGTTTTCCTTAGGCTGTACACACA	2029
Db	864	TCCTTTGGAGGTTTACAGAAATAGTAGACTGGCAGAAATAATGAAATCTTCAATCCAAG	923	1941	CTAGTGA---CCAAATGCCACAACAAATGAAATAATGTTTGTGGGTTTCTTACAACT	1997
QY	950	GACGATCATTTTAAAGTCCCTGCATCTACTGCTGTGTTCTCATGACACTCAAAAC	1009	2030	GTGAATGGAATTTGGAAGATGGACTCAAGGAACAAGGCCGTGATGCTGCGGCTATCTT	2089
Db	924	GATGATCATTTCTCAGCTCTCCGGCATCTCAGCGGCTGTATTCATGCGTACAGGAAAC	983	1998	TTTAATGATATAGCAAAAGAGGAGCTGAGAGCAAGGCCGATGCTAGGCTACATT	2057
QY	1010	AAACGATGCTCCACTTTCTCAACTTCGTGCTCAGCAAAATTTGGGACTACGTTCTTTCG	1069	2090	CGMAAGTATCGGAGGCTTCTCGCATCTATACCAAGAACCGAATGTTGCGGAGCA	2149
Db	984	AAAGTGTGTTGATTTCTTGAACCTTCTGAGTCTTCTGAGAAATTCGGAACCATGCTGCTGT	1043	2058	CAAAATGTTTGGAAAGTCCAATCTTGAAGCTTACACGAAGAACAGAAATGGTCTGAAGCT	2117
QY	1070	CATTACCCACTTGATCTATTGAAACCTCTCGGCTGTCGATACAGTTGAAACGCTTGGGA	1129	2150	AAGTATGTCGCGACATTTCAACGAATATGTGGAATATGGAATGCAAGTCTCCATAGACTTGG	2209
Db	1044	CATATCCGCTGATCTATTGAAACGTTTGTGGCGGTTGATACAGTGTAGGCGGCTAGGT	1103	2118	AAATATGTGCCATCTCTCAATGAATACATAGAAATGCGAGTGTCTCAATAGCATTTGGA	2177
QY	1130	ATCGATGCTATTTCAAGAAAGAAATCAAGAAATCTCTGATATCAATGTTTATAGTACTGG	1189	2210	ACAGTCTGACTAAACTCAATCTTTTCTGAGAAATTTACTTCTCTGATTTACATTTTACAG	2269
Db	1104	ATCGATGCTATTTCAAGAGGAGATCAAGAGCATTTGATATGTTTACAGCCATTTGG	1163	2178	ACAGTGGTCTCTCATTTAGTGTCTTTTCTGAGGAGTCTTACAGATGAAGTACTCTCC	2237
QY	1190	GACGCGAAGAGCGGTGGATGGGCAAGATGCAATCTTCTGATGTGATGACAT	1249	2270	CAAGTAGACCTTCGGTCCAAATTTCTGCAATCTTGTGCTTTTCTGAGTGGAGCACTAATCAAT	2329
Db	1164	GA---CGMAAGAGCATTTGATGGGCGAGAGAAATCTGTTCTGATATTTGATGATACA	1220	2238	AAAATTTGATCGGCAATCTAGATTTCTTCAACTCATGGGCTTAAACAGGGCGTTTGGTGAAT	2297
QY	1250	GCCATGGCTTTAGATCTCTGAGACTTCATGATACAAATGATCTTCTGATGTTCTGGAG	1309	2330	GAACCAAGACTTACCGGCCGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTACAGTGC	2389
Db	1221	GCCATGGGCTTCTGAAATCTTTGAGATTTACATGGATACATGATATCTCTCAAGT	1280	2298	GAACCAAACTTATCAGGACAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATACAATGT	2357
QY	1310	AATTTTCAAGACGAGAAAGAGACTTCTTTTGTCTTTGCGGTCGCAAGCAAAATTTGGTGTG	1369	2390	TACATGAGGGAATATCCGAGTGCACAGAGGAAGAGCTCTAAGTCTATGTTTATGATC	2449
Db	1281	ACATTTAGATGAGATGGAGTCTTTTGTCTTTGCTTCTTGGGTCAAAACACAGAGAGGAT	1340	2358	TATATGAAGGACCATCTCTAAATCTCTGAAGAAAGAGCTCTACAACTGTCTATAGTGC	2417
QY	1370	ACGATATCTTAACTTTATAGTGTCAAGATGTTTTCGGGAGAAAGATTAATG	1429	2450	ATCGACACGCACTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCA	2509
Db	1341	ACAGACATGTTAAACGTCATCGTGTTCATGTTTTCATTTTCGGAGAAACGATCATG	1400	2418	ATGGAATATGCCCTCGAAGAGTTGAATAGGAGTTTGTGNA-----TAACAAATACCG	2471
QY	1430	GAAGAGCTAAGACTTCACTACAAATCATCTCCAAATGCTCTTGCACAAACCAAGCA	1489	2510	TTGTGTGTCAGAAAGACTGCTGTTTCAACACTGCAAGAGTGTGATGAGTGTGTTTATGATC	2569
Db	1401	GAAGAGCAAACTCTGTACCGAAAGTATCTGAGGAATGCTCTGGAATAATGTGGATGCC	1460	2472	GATATTTACAAAGACTGTTTGTGAACTGCAAGAAATAATGCAACTCTTTTATATGCAA	2531
QY	1490	TTTGATAGTGGGCTGTCAAGAGGATCTTCTGAGAGGTTGGATGATGCTATTAAGTAT	1549	2570	AGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTGACGCCGAACTCTT	2626
Db	1461	TTTGACAAATGGGCTTTTAAAGAAATTTTCGGGAGGTTAGATGATGCACTCAAAATAT	1520	2532	GGGATGGTTTGAACACTTATCAGATGATGGAATTAAGAGCATGTCAAAAATTTGCTC	2591
QY	1550	CCGTGGCATAGAGTATGCCAAGATTTGAGGCAAGAGTTTACATAGAGCAATTTGGATCA	1609	2627	TTCGATCTCTGGCTAGCATTA	2648
Db	1521	CCCTGGCATAGAGTATGCCAAGGTTTGGAGCTAGAGCTATATTGAAACTATGGGCCA	1580	2592	TTCCAACCAAGTTGCTAGATTA	2613
QY	1610	AATGATGCTGGCTGGGAGACTGTGTATTAAGATGCTATATGTGAGCAACGAAATAT	1669			
Db	1581	GATGATGCTGGCTGGGAAACTGTATATATGATGCCATACATTTTCAATGAAAGTAT	1640			

RESULT 2  
US-09-887-586A-55  
; Sequence 55, Application US/09887586A  
; Patent No. 6495354



GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 64953541, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/887,586A  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(2606)  
; OTHER INFORMATION: abietadiene synthase  
US-09-887-586A-55

Query Match 40.3%; Score 1091.4; DB 4; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

Qy	350	GTGCATGAGACAATCAATAAAGAAATCCAGACATGTTGGTGAAGAAATCCAGTGCATGTTT	409
Db	324	GTTGCAGCATCAGACGAGAGCGTATCGAGACATTAATATCCGAGATTAAAGATATGTTT	383
Qy	410	CAGTCCATGGCGACGTTGAACCAATCCATCTGCATATGATACAGCTTGGTGGCAAGA	469
Db	384	AGATGTTATGGGCTATGGGAAACGAATCCCTCTGCATATGACATCTGCTTGGGTAGCAAG	443
Qy	470	ATTCCGCTCAATTTGACGGCTCTGGTGACCCCAATTTCCCAACCGTTCAATGATTTCTG	529
Db	444	ATTCCAGCAGTTGATGGCTCTGACAAACCTCACTTCTCTGAGACGGTTGAATGATTTCTT	503
Qy	530	AACAAATCAATCCAGATGGCTCTGGGGTAGAGATGCAATTTTCTGGGCTATGACAGA	589
Db	504	CAAAATCAGTTTGAAGATGGGTCTTGGGGTGAAGGATTTCTACTTCTTGGCATATGACA	563
Qy	590	GTTTTTAAACACTCTCGCTGCTCTCTCACTCTCAAAATATGGAATAAGGGCGACATTCAA	649
Db	564	ATACTGGCTACATTCGATGATTAATTAACCTTACCCTCTGGGCTACTGGGGAGACAAA	623
Qy	650	GTGCAGAAAGGGTTGAGTTTGTGAGAAACACATGGAAGAAATGAAGGACGAAGCTGAC	709
Db	624	GTACAGAAAGGTTGAATTTCTTCAGGACACAAGCTGGAAGATGGAAGATGAGCTGAT	683
Qy	710	AATCAGAGGCCAAGTGAATTCGAGGTCGTGTTTCTGCAATGTTAGATGAAGCAAAAGC	769
Db	684	AGTCATAGGCCAAGTGAATTTGAATAAGTATTTCTGCAATGCTTAAAGGAAGCTTAAATC	743
Qy	770	TTGGGATGGATCTTCTTATCACTCCCTTTTCACTCCCAATCCCAAAATCCCAAAAGGCGAG	829
Db	744	TTAGGCTTGGATCTGCCCTTACGATTTGCAATCTCTGAAACAAATCATCGAAAGCGGGAG	803
Qy	830	AAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCAATCAACCATCAGACGGCTGCTCTAC	889
Db	804	GCTAAGCTTAAAGAGATTTCCCACTGATGTTCTTATGCCCTTCCCAACAACTTATGAT	863
Qy	890	TCTCTGAGGGTTTTCAGAGATGTGGTGAATGGGCAAGAGATCAAAATCTTCAATCAAGA	949
Db	864	TCTTTGGAAGTTTACAGAAATAGTAGACTGGCAGAAATAATGAAACTTCAATCCAAG	923
Qy	950	GACGATCAATTTTAAAGTCCCTGCACTACTGTTGTTCTTTCATGACACACTCAAAAC	1009

Db	924	GATGGATCAATTTCTCAGCTCTCGGCATCTACAGCGCTGTATTTATCATGCGGTACAGGGAAC	983
Qy	1010	AAACGATGCTCCACTTTCTCAACTTCGTCCTAGCAAAATTTGGCGACTAGTTCCTTGC	1059
Db	984	AAAAAGTCTGGATTTCTTGAACCTTTGTCTTGAAGAAATTCGAAACCATGTGCTTGT	1043
Qy	1070	CATTACCCACTTCATCTATTGTAACGCTCTGGGCTCTGCATACAGTGTGAACGCTTGGGA	1129
Db	1044	CATATCCGCTTGATCTATTGTAACGTTTGGGCGGTTGATACAGTGTGAGCGGTAGGT	1103
Qy	1130	ATCGATCGCTATTTCGAAGAAAGAAATCAAAAGAAATCTCTGGATTAAGTTCAGTACTGG	1189
Db	1104	ATCGATCGCTATTTCGAAGAGGAGATCAAGGAAGCATTTGGATTATGTTTACAGCCATTGG	1163
Qy	1190	GACGCCAAGAGCGGTGGGATGGGCAAGATGCATCTTATCTCTGATGTCGATGACACT	1249
Db	1164	GA---CGAAAGAGCATTTGGATGGGCGAGAGAAATCTCTGTTCTGTATTTGATGATACA	1220
Qy	1250	GCCATGGTCTTAGAATCTCGAGACTTTCATGGATACAAATGTATCTTCAGATGTTCTGGAG	1309
Db	1221	GCCATGGGCTTCCGATCTTGAGATTACATGGATACAAATGTATCTTCAGATGTTTAAAA	1280
Qy	1310	AATTTACAGACGAGAAAGGAGACTTCTTTTGTCTTCCCGCTCAAAACGCAAAATTTGGTGT	1369
Db	1281	ACATTTAGATGAGATGGGAGTCTTTTGTCTTCTTGGTCAAAACAGAGAGAGTT	1340
Qy	1370	ACCGAATCTTAACTTTATAGATGTTCAAGATATGTTTTCGGGAGAAAGATATG	1429
Db	1341	ACAGACATGTTAAACGTCATCGTTGTTCATGTTTCAATTTCCGGGAGAAACGATCATG	1400
Qy	1430	GAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGCACCAAAACACGCA	1489
Db	1401	GAAGAGCAAAACTCTGTACCGAAAGGTATCTGAGGAATGCTCTGGAATAATGTGGATGCC	1460
Qy	1490	TTTGATAAGTGGGCTGTCAAGAGGATCTTCTCGAGAGGTGAGATGCTATATAAAGTAT	1549
Db	1461	TTTGACAAATGGGCTTTTAAAGAGATATTCGGGAGAGGTAGATGATGCACTCAATAT	1520
Qy	1550	CCGTGGCAATAGATATGCAAGATTGGAGCAAGAGTTACATAGAGCAATTTGGATCA	1609
Db	1521	CCCTGGCAATAGATATGCAAGGTGGAGCTAGAAAGCTATATTTGAAACATATGGGCA	1580
Qy	1610	AATGATGCTGCTGGTGGGAGAGCTGTATAGATGCTATATGAGCAACGAAATAT	1669
Db	1581	GATGATGTTGGCTTGGAAACCTGTATATATGATGCCATACATTTGGAATGAAAGTAT	1640
Qy	1670	TTGAGCTGCGCCAAATTTGGAATTTCAATATGCTGAGCGCTTACACCAAAAGAGACTCAA	1729
Db	1641	TTAGAACTAGGAAACTGGACTTCAATAAGGTGCACTTATACACCAAAACAGAGCTCAA	1700
Qy	1730	CACATGTCAGCTGGTGGAGAGAAATCGGATTCATATGATCTTACATTCACCCCGCAGCG	1789
Db	1701	GATCTTCAAGGTGGTGGAAATCATCCGGTTTACCGGATCTGAATTTCACTCGTGAGCGT	1760
Qy	1790	CTGTGGAATGATTTTCTCAGTGGCGGTTAGTATGTTTGGCCAGAAATTCGCTGCTGT	1849
Db	1761	GTGACGGAATAATTTCTCAACCGGCTCTTATCTTGGAGCCGAGTTTCTTAAGTGC	1820
Qy	1850	AGAATTCCTTATGCAAGACTTCTTGCTCGCAGTATTTAGACAGTCTTTAGACACACC	1909
Db	1821	AGAGAGTTTATACAAAACTTCCAAATTTCACTGTTATTTAGATGATCTTTATGACGCC	1880
Qy	1910	CACGATCTCTGGATGATCTTAAATGTTTCTTGAACGGTCCGAAGATGGGATATCTCT	1969
Db	1881	CATGGATCTTTAGACGATCTTAAGTCTTACAGAAATCAGTCAAAAGATGGGATCTATCA	1940
Qy	1970	GTGCTGATAGCGTTCCGGATATACAGTTGAAAGTTTGTCTTCTAGGCTGTACACACA	2029
Db	1941	CTAGTGA---CCAAATGCCCAACAAATGAAATATGTTTGTGGGTTTCTACATACT	1997
Qy	2030	GTGAATGATTTGGAAGAGATGGAATCAAGGCAACAGCGCTGATGTGCTGGGCTATCTT	2089
Db	1998	TTTAATGATATAGCAAAAGAGGACGTGAGAGGCGAGCGGATGTGCTAGGCTACATTT	2057

2090 CGAAAAGTATGGAGGGTCTCTCGCATCGTATACCAAGAGCGCGATGGTGGCAGCA 2149  
Db CAAAATGTTTGGAAAGTCCAACTTGAAGCTTTACAGAAAGAGCAGAAATGGTCTGAAGCT 2117  
Qy AAGTATGTGGCGACATTCACAGAAATATGTGCAAAATGCAAAAGTGTCCATAGCACTTGGC 2209  
Db AAATATGTGCCATCTTCAATGAATACATAGAAATGGAGTGTGTCAATAGCATGGGA 2177  
Qy ACAGTCGTAATAACTCTTTTTCATCTGAGAAATTAATCTTCTGATTAATTAATTAACAG 2269  
Db ACAGTCGTTCTCATTAAGTGTCTTTTTCATCTGAGAAATTAATCTTCTGATTAATTAACAG 2237  
Qy CAAGTAGACCTTCGGTCCAAATTTCTGATCTTGTGCTTTTGAATGAGAGCTAATCAAT 2329  
Db AAAATTTGATCGGAATCTAGATTTCTTCACTCATCTGAGGCTTAAAGGGGTTTGTGAT 2297  
Qy GACACCAAGACTTACAGGCGGAGAGAAACCGTGTGAATTTGGTTTCCAGGCTACAGTGC 2389  
Db GACACCAAACTTATACAGGCGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATACAAATGT 2357  
Qy TACATGAGGAAATTCGGAGTGCACAGAGGAAGCTCTAAGTCAATGTTTATGTTATC 2449  
Db TATATGAAGGACCATCTTAAATCTCTGAAGAAAGAGCTCTACAAATGTCTATAGTGT 2417  
Qy ATCGACAGGCACTGAAGAAATTTGAATTTGGGAGTTGGCCAACTCCAGCGAGCAATGCCCA 2509  
Db ATGGAAATGCCCCGAGAGTTGAATTTGAATTTGGGAGTTGGCCAACTCCAGCGAGCAATGCCCA 2471  
Qy TTGTGTGTGAGAGAGTCTGTCTCAACTCTCAAGTGTGAATTTGGTTTCCAGGCTACAGTGC 2569  
Db GATATTTACAAAGACTGTGTTTGAATTTGAATTTGGGAGTTGGCCAACTCCAGCGAGCAATGCCCA 2531  
Qy AGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGAGCATGTACAGCGAACTCTT 2626  
Db GGGGATGGTTTGACATCTATCAGATGATGGAATTTAAAGAGCATGTCAAAATTTGCCCTC 2591  
Qy TTCTGATCTCTGGCGTAGCATA 2648  
Db TTCCAACAGTTGGCTAGATTA 2613

RESULT 3  
US-09-895-752-55  
; Sequence 55, Application US/09895752  
; Patent No. 6559297  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6559297, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/895, 752  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(2606)  
; OTHER INFORMATION: abietadiene synthase

US-09-895-752-55

Query Match 40.3%; Score 1091.4; DB 4; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;  
Qy 350 GTCGATGAGACAATCAATAAAGAAATCCAGACATTTGGTGAAGGAATCCAGTGCATGTTT 409  
Db GTTCGAGCATCAGACGAGAGCGGTATCGAGACATTAATATCCGAGATTAAGAAATATGTTT 383  
Qy 410 CAGTCCATGGCGAGCGGTGAACGAATCCATCTGCATATGATACAGCTTTGGGTGGCAAGA 469  
Db AGATGTATGGGCTATGGCGAAACGAATCCCTCTGCATATGACATCTGCTTGGGTAGCAAGG 443  
Qy 470 ATTCGGTCAATTTGACGGCTCTGGTGCACCCCAATTTCCCAAAACGCTTCAATGGATTTCTG 529  
Db ATTCAGCAGTTGATGGCTCTGACAAACCTCACCTTCTGAGACGGTGAATGGATTTCTT 503  
Qy 530 AACAAATCACTGCCAGATGGCTCTGGGTGAGGAGTGCATTTTCTGCGGTATGACAGA 589  
Db CAAAATCAGTTGAAGATGGGTCTTGGGGTGAAGGATTTCTACTTCTTGGCATATGACAGA 563  
Qy 590 GTTTTAAACACTCTCGCTGCTCTCTCACTCTCAAAATATGGAATTAAGGGCGACATTTCAA 649  
Db ATACTGGCTACACTTGCATGATTAATTAACCTTACCTCTGCGGTACTGGGGAGACACAA 623  
Qy 650 GTGCAGAAAGGGTTGAGTTTGTGAGAAAACACATGGAAGAAATGAAGACGAAGCTGAC 709  
Db GTACAGAAAGGTATTGAATTTCTTCAAGACACAAGCTGGAAAGATGGAAGATGAAGCTGAT 683  
Qy 710 AATCACAGCCCAAGTGAATTCGAGGTCTGTTCTGCAATGTTAGATGAAGCAAAAGC 769  
Db AGTCATAGCCCAAGTGGATTTGAAATAGTATTTCTGCAATGCTAAAGAAAGCTAAATC 743  
Qy 770 TTGGGATTTGATTTCTTCTTATCACCTCTTCTTCACTTCCAAATCCCAAAATCCCAAAAGCCCGAG 829  
Db TTAGGCTTTGATCTGCTTACGATTTTGCATTTCTTCAAAACAAATCATCGAAAGCGGGAG 803  
Qy 830 AAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCAATCAACCATCAGACGCGTGTCTCTAC 889  
Db GCTAAGCTTTAAAGGATTTCCCACTGATGTTCTTATGCCCCCTTCCAAACAGCTTATTTGAT 863  
Qy 890 TCTCTGGAGGTTTGAAGATGTGGTGAAGTGGCGACTGGCAAGAGATCAAAATCTTCAATCAAGA 949  
Db TCTTTGGAAGGTTTACAGAAATAGTAGACTGGCAGAAATTAATGAATCTTCAATCAAG 923  
Qy 950 GACGATCATTTTAAAGCTCCCTGATCTACTGTTGTCTTCTTCAATGACACTCAAAAC 1009  
Db GATGGATCATTTCTCAGCTCTCGGCATCTACAGCGGCTGTATTTCAATGCGTACAGGAAC 983  
Qy 1010 AACGATGCTCCACTTTCTCAACTTCTGCTGCTCAGCAAAATTTGGCGACTACGTTCCCTTGC 1069  
Db AAAAGTGTCTTGAATTTCTTGAACCTTTGCTTGAAGAAATTCGAAACCATGTGCTTGT 1043  
Qy 1070 CATTACCCACTTGTATCTATTGAAACGCTCTGGGCTGTGATACAGTGTGAACGCTTGGGA 1129  
Db CACTATCCGCTTGTATCTATTGAAACGCTTGTGGGCGGTTGATACAGTGTGAGCGGCTAGGT 1103  
Qy 1130 ATCGATGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGATTTACGTTTATAGTACTGG 1189  
Db ATCGATGCTTATTTCAAGAAAGAAATCAAGAAAGCAATTTGAATTTATGTTTACGCCATTGG 1163  
Qy 1190 GACCGCGAAAGAGCGGTGGGATGGGCAAGATGCAATCTTATCTGATGTGATGACACT 1249  
Db GA---CGAAAGAGGCAATTTGATGGCGGAGAGAGATCTCTGCTGATATTTGATGATACA 1220  
Qy 1250 GCCATGGGCTTTAGAAATCTGAGACTTCTGATGATACATGATATCTTCAAGATTTCTGGAG 1309  
Db GCCATGGGCTTTGAAATCTTGAATTTACATGATGATGATGATGATGATGATGATGATGAT 1280  
Qy 1310 AATTTGAGACGAGAGAGAGACTTCTTTTGTCTTTCGCGTCAAAAGCAAAATTTGGTGTG 1369  
Db ACATTTAGATGAGATGGGAGTCTTTTGTCTTCTTGTCTTCTTGTCTTCTTGTCTTCTTGTCT 1340

```
Qy 1370 ACCGATATCTTACCTTTATAGATGTTCAAGTATGTTTTCGGGAGAAAGATATG 1429
Db 1341 ACAGACATGTTAAACGTCATCGTTGTTTCATGTTTTCCTCGGAGAAACATCATG 1400
Qy 1430 GAAGAAGCTAAGACCTTTCACCTACAAATCATCTCCAAATGCTCTTGCCAAATAACAAGCA 1489
Db 1401 GAAGAAGCAAACTCTGTACCGAAAGGTATCTGAGGAATGCTCTGGAATATGGATGCC 1460
Qy 1490 TTTGATTAAGTGGGCTGTCAAGAAGATCTTCTCGGAGAGTGGAGTATGCTATTAAGTAT 1549
Db 1461 TTTGACAAATGGGCTTTTAAAAAGAAATATTTCGGGAGAGGTAGAGTATGCACTCAAAATAT 1520
Qy 1550 CCGTGGCATAGAGTATGCCAAGATTTGGAGCAGAAGTTTACATAGAGCAATTTGGATCA 1609
Db 1521 CCCTGGCATAGAGTATGCCAAGGTTGGAGGCTTGAAGCTATATTGAAAACATATGGGCCA 1580
Qy 1610 AATGATGCTCTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTGAGCAACGAAAATAT 1669
Db 1581 GATGATGCTGGCTTGGAAAACCTGTATATATATGATGCCATACATTTGCAATGAAGAATAT 1640
Qy 1670 TTGAGCTGGCCAAATTTGGACTTCAATATATGCTGAGGCCCTTTACACCAAAAAGGAGACTCAA 1729
Db 1641 TTAGAACTAGCGAACTGGACTTCAATAAGGTGCGTCTATACACCAAAACAGAGCTTCAA 1700
Qy 1730 CACATTTGCTAGCTGGTGGAGAGAAATCGGATTTCAATGATCTTACATTTACCCGCCAGCG 1789
Db 1701 GATCTTCCAAAGGTGGGAATCATCCGGTTTCAAGGATCTGAATTTCACTCGTGAGCGT 1760
Qy 1790 CCTGTGAAATGATTTCTCAGTGGCGGTAGTATGTTTGAGCCAGAAATTCGCTGCTTGT 1849
Db 1761 GTGACGGAATATATTTCTCACCGGATCTTTATCTTTGAGCCGAGTTTCTTAAGTGC 1820
Qy 1850 AGAATTCCTATGCAAGACTTCTTGCTCGCAGTATTTCTAGACGATCTTTAGCACACC 1909
Db 1821 AGAGAGGTTTATACAAAACCTTCCAAATTCACCTGTTATTTAGATGATCTTTATGACGCC 1880
Qy 1910 CACGGATCTCTGGATGATCTTAAATTTGTTCTCTGAAGCGGTCCGAAGATGGATATCTCT 1969
Db 1881 CATGGATCTTTAGACGATCTTAAGTTGTTTACAGAAATCAGTCAAAAGATGGGATCTATCA 1940
Qy 1970 GTGCTGATAGCGTTCCGGATTAATCAGTTGAAGTTTTCCTTAGGCGCTGTAACACA 2029
Db 1941 CTAGTGA---CCAAATGCCAAACAAATGAAATATGTTTGTGGGTTTCTACAACTACT 1997
Qy 2030 GTCAATGGATTTGAAAGATGGAATCAAGSAAACAGCCCGTGTATGTCTGGGCTATCTTT 2089
Db 1998 TTTAATGATATAGCAAAAGAGGACGTGAGAGGCAAGGCGCGATGTGCTAGGCTACATTT 2057
Qy 2090 CGAAAGTATGGGAGGCTTGTCTCGCATCGTATACCAAGAGCCGAATGGTGGCAGCA 2149
Db 2058 CAAAATGTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGAGCAAGATGCTCTGAAGCT 2117
Qy 2150 AAGTATGTCGCGACATTTCAAGATATGTGGAATATGCCAAAGTGTCCATAGCACTTGG 2209
Db 2118 AAATATGTCGCATCTTCAATGAATAATAGAGAAATCGAGTGTGTCAATAGCAATGGGA 2177
Qy 2210 ACAGTCTGTACTAACTCAATCTTTTTCACCTGGAGAAATTAATCTTCTGATTAATTTTACAG 2269
Db 2178 ACAGTCTGTCTAATAGTGTCTTTTTCACCTGGGAGGTTCTTACAGATGAAGTACTCTCC 2237
Qy 2270 CAAGTAGACCTTCGGTCCAAATTTCTGCACTTGTGTCTTTGATCTGGAAGCACTTAATCAAT 2329
Db 2238 AAAATTGATCGCGAATCTAGATTTCTTCAACTATGCGGCTTAAACAGGCGGTTTGGTGAAT 2297
Qy 2330 GACACCAAGACTTACAGGCGGAGAGAAACGTTGTAATTTGGTTTCCAGCGTACAGTGC 2389
Db 2298 GACACCAAACTTATCAGGACAGAGAGGTTCAAGTGGAGTGGCTTCTGCCATACAATGT 2357
Qy 2390 TACATAGGGGAAATCCGGAGTGCACAGAGAAAGACTCTTAAGTCAATGTTTATGGTATC 2449
Db 2358 TATATGAGGACCATCTTAAATCTCTGGAAGAGAGCTCTTACAACTATGCTATAGTGC 2417
```

```
Qy 2450 ATCGACAAACGCACTGAGGAATTTGAATTTGGAGTTTGGCCAAACCCAGCGACGATGCCCA 2509
Db 2418 ATGGAAAATCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAAAATACCG 2471
Qy 2510 TTGTGTGTGAGAAAGACTGCTGTGTTCAACACTGCAAGAGTGAATGAGCTGTTTATATATGAC 2569
Db 2472 GATATTACAAAGAGACTGTTTGTGAACTGCAAGATAATGCAACTCTTTTATATGCAA 2531
Qy 2570 AGAGATGGCTTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTFCAGCCGAATCTTT 2626
Db 2532 GGGGATGGTTTGACACTATCATCATGATATGGAATTAAGAGCATGTCAAAAAATGGCTC 2591
Qy 2627 TTCGATCCTGTGGCGTAGCATTA 2648
Db 2592 TTCCAAACAGTTGCCTAGATTA 2613

RESULT 4
US-09-903-012B-55
; Sequence 55, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65696561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 2861
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2606)
; OTHER INFORMATION: abietadiene synthase
US-09-903-012B-55

Query Match 40.3%; Score 1091.4; DB 4; Length 2861;
Best Local Similarity 68.4%; Pred. No. 0;
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

Qy 350 GTCGATGAGCAATCAATAAAGAAATCCAGACATTTGGTGAAGGAAATCCAGTGCATGTTT 409
Db 324 GTTGCAGCATCAGACGAGAAAGCGTATCGAGCAATTAATATCCGAGATTAGAATATGTTT 383
Qy 410 CAGTCCATGGCGGACGGTGAACCAATCCATCTGCATATGATACAGCTTCGGTGGCAAGA 469
Db 384 AGATGATGGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTTGGGTAGCAAG 443
Qy 470 ATTCCGTCAATTGACGGCTCTGTGACCCCAATTTCCCAACCGTTCAATGGATTTCTG 529
Db 444 ATTCCAGCAGTTGATGGCTCTGCAACCCCTCACTTTTCTGAGACGGTTGAATGGATTTCTT 503
Qy 530 AACAAATCAATCGCAGATGGCTCGTGGGGTGAAGAGTGCATTTTCTGGCGTATGACAGA 589
Db 504 CAAATCAGTTGAAGATGGTCTTGGGTGAAGGATTTCTACTTCTTGGCATATGACAGA 563
Qy 590 GTTTTAAACACTCTCGCCTGCTCTCTCACTCTCAAAATATGGAATTAAGGGCGCAATTCAA 649
Db 564 ATACTGGCTACACTTGCATGTAATTATTACCTTACCCTCTGGCGTACTTGGGGAGACAA 623
Qy 650 GTCGAGAAAGGGTTGATTTGTGAGAAAACACATGGAAGAAATGAAGGAGCAAGCTGAC 709
```

Db	624	GTACAGAAAGGTAATGAAATCTTCAGGACACAAGCTGGAAAGATGAAAGCTGAT	1701	GATCTTCGAAAGGTGGTGGAAATCATCGGTTTTCACGGATCTGAAATTTCACTCGTAGCGT	1760
Qy	710	AATCACAGCCCAAGTGGATTCGAGTCTGTTTCTGCAATGTTAGATGAAGCAAAAAGC	1790	CTGTGGAAATGTAATTTCTCAGTCGCGGTAGTATGTTTGGAGCCAGAAATTCGCTCTTGT	1849
Db	694	AGTCATAGCCCAAGTGGATTTGAATAGTATTTCTGCAATGCTAAAGGAGCTAAATC	1761	GTGACGGAATATATTTCTCACCGCATCTTTATCTTTGAGCCGAGTTTCTAAGTGC	1820
Qy	770	TTGGGATTTGGATCTTTCCTTATCATCCTCCCTTTCATCTCCCAATCCACAAAGCCGAG	1850	AGAAATGCTCTATGCAAGACCTTCTTGCTCGCATTTCTTAGACGATCTTTTACGACACC	1909
Db	744	TTAGGCTTTGGATCTGCTTACGATTTGGCAATCTCTGAAACAAATCATCGAAAGCGGAG	1821	AGAGAGTTTATACAAAACCTTCCAAATTTCACTGTTATTTTAGATGATCTTTATGACGCC	1880
Qy	830	AAAAGCTTCAAAAGATCCCTCAATGTTCTTCAATACCATCACAGCGGTGCTCTAC	1910	CACGATCTCTGGATGATCTTAAATTTGTTCTCTAGAGCGGTCCGAGATGGATATCTCT	1969
Db	804	GCTAAGCTTAAAGGATTTCCCACTGATGTTCTATGCGCTTCCAAACAGCTTATGTAT	1881	CATGGATCTTTAGACGATCTTAAAGTTGTTTCAAGATCAGTCAAAAGATGGATATATCA	1940
Qy	890	TCTCTGGAGGGTTTCCAAAGATGTGTGGACTGGCAAGAGATCACAAATCTTCAATCAAGA	1970	GTGCTGGATAGCTTTCCGGGATAATCAGTTGAAAGTTTCTCTCTAGGGCTGTACAAACA	2029
Db	864	TCTTTTGGAGGTTTACAGAAATAGTAGACTGGCAGAAAATAATGAAACTTCAATCCAA	1941	CTAGTGGH---CCAAATGCCACACAAATGAAATATGTTTGTGGGTTTCTACAATACT	1997
Qy	950	GACGATCAATTTTAAAGCTCCCTGCACTCTGCTTGTGTTCTTATGCACTCAAAAC	2030	GTGAATGGATTTGGAAAAGATGGACTCAAGGAAACGAGCCGCTGATGCTGGCTATCTT	2089
Db	924	GATGATCATTTCTCAGCTCTCCGCACTACAGCGGCTGTATTCATGCTACAGGGAAC	1998	TTTAATGATATAGCAAAAGAGGACGCTGAGAGGCAAGGGCGCGATGTGTAGGCTACAT	2057
Qy	1010	AAACGATGCTCCACTTCTCAACTTCGFGCTCAGCAAAATTTGGGACTTACGTTCTCTTC	2090	CGAAAAGTATGGAGGGCTTCTCGCATCTGATATACAAAGAGCCGAAATGGTCCGACGA	2149
Db	984	AAAAAGTCTTGATTTCTTGAACCTTGTCTTGAAGAAATTCGAAACCATGTGCTTGT	2058	CAAAATGTTTGGAAAGTCCAACTTGAAGCTTACAGAAAGACAGAAATGGTCTGAAGCT	2117
Qy	1070	CAATACCACTTGAATATTTGAAGCCCTCTGGGCTGTCGATACAGTTGAAACGCTTGG	2150	AAGTATGTGCCGACATTCACAGAAATATGTGGAATATGTGGAATATGCAAGTGTCCAT	2209
Db	1044	CACATCCGCTTGATCTAATTTGAACGTTTGTGGGGGTTGATACAGTTGAGCGGCTAG	2118	AAATATGTGCCATCTTCAATGAAATACATAGAAATGCGAGTGTGTCAATAGCATTTGG	2177
Qy	1130	ATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGATTTAGCTTATAGTACTGG	2210	ACAGTCGTACTAACTCAATCTTTTCTCAGTGGAAATTTCTCTGATTTACATTTTACAG	2269
Db	1104	ATCGATCGCTATTTCAAGAGGAGATCAAGAGCAATTCGATTAATGTTTACGCCATGG	2178	ACAGTCGTCTCATTTAGTGTCTCTTTTCTGAGGAGGTTCTTACAGATGAGTACTCTCC	2237
Qy	1190	GACGCCGAAGAGCGGTGGGATGGGCAAGATGCAATCTTATCTGATGTCGATGACACT	2270	CAAGTAGACCTTTCGGTCCAAATTTCTGCAATCTTGTGTTCTTTGACTGGAGCACTTAAT	2329
Db	1164	GA---CGAAGAGGCATTTGATGGGCGAGAGAGAAATCTCTGATTTGATGATACA	2238	AAATTTGATCGGAATCTAGATTTCTTCACTCATGGGCTTAACAGGGCGTTTGGTGAAT	2297
Qy	1250	GCCATGGGCTTTAGAAATCTGAGACTTCAATGATGATGATGATGATGATGATGATG	2330	GACACCAAGACTTACAGGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGC	2389
Db	1221	GCCATGGGCTTTGAAATCTTGAATTTACATGATGATGATGATGATGATGATGATG	2298	GACACCAAACTTATCAGGCAGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATACAATGT	2357
Qy	1310	AAATTCAGAGACGAGAAAGAGACTTCTTTTGTCTTTCGCGTCAAAACGCAAAATTTGTG	2390	TACATGAGGAAATTCGCGAGTGCACAGAGGAGAGAGCTTAAGTCAATGTTTATGTTATC	2449
Db	1281	ACATTTAGAGATGAGAAATGGGAGTTCTTTTGTCTTCTGGGTCAAAACACAGAGAGG	2358	TATATGAGAGCAATCTTAAATCTCTGAAGAAAGAGCTTACAAACATGCTATAGTGC	2417
Qy	1370	ACCGATAATCTTAACTTTATAGATGTTCAAAAGTATGTTTCCGGGAGAAAGATAATG	2450	ATCGCAACGCACTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCA	2509
Db	1341	ACAGACATGTTAAACGCTCAATCGTTGTTTCAATGTTTCAATTTCCGGGAGAAACGATC	2418	ATGGAATATGCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAAACAAATACCG	2471
Qy	1430	GAAGAGCTAAGACCTTCACTCAAAATCTCTCAAAATCTCTTGGCAAAACCAACGCA	2510	TTGTGTGAGAGAGACTGCTGTTCAACACTGCAAGAGTATGATGAGCTGTTTATATGTAC	2569
Db	1401	GAAGAGCAAAACTCTGTACCGAAAGGATCTGAGGATGCTCTGGAATTTGGATGCC	2472	GATATTTACAAAGAGCTGTTTGTGAACCTGCAAGAAATATGCAACTCTTTTATATGCAA	2531
Qy	1490	TTTGATAAGTGGGCTGTCAGAAAGGATCTTCTGGAGAGTGGAGTATGCTATAAGATAT	2570	AGAGATGGCTTTTGGTATCTC---TGACAAAGAGATGAAAGACCATGTACGCCGAACTCTT	2626
Db	1461	TTTGACAAATGGGCTTTTAAAGAAATATTTCCGGGAGGATGATGATGCACTCAAAAT	2532	GGGATGTTTGGACATCTACATGATATGGAATTTAAAGAGCATGTCAAAATTTGCCTC	2591
Qy	1550	CCGTGGCATAGATGTCAGAAATTTGGAGGCAAGATTTACATAGAGCAATTTGGATCA	2627	TTCCGATCTTGTGGCGTAGCATA	2648
Db	1521	CCCTGGCATAGAGTATGCCAAGGTTGGAGGCTTGAAGCTATATTTGAAACTATGGGCCA	2592	TTCCAAACGAGTTGCCCTAGATTA	2613
Qy	1610	AATGATCTCTGGCTGGGAGAGACTGTGTATAAGATGCTATATGTGAGCAACGAAATAT			
Db	1581	GATGATGTGGCTTGGAAACCTGTATATATGATGCTATATGATGATGATGATGATGAT			
Qy	1670	TTGAGCTGGCAAAATTTGAACTTCAATATGTTGAGGCTTACACAAAGAGGACTCAA			
Db	1641	TTAGAACTAGCGAACTGGACTTCAATAAGGTGAGTCTATACACCAACAGAGCTTCAA			
Qy	1730	CACATTTGCTGGTGGAGAGATCGGATTTCAATGATCTTACATTTACCCCGCAGCGG			

RESULT 5  
US-09-900-797-55  
; Sequence 55, Application US/09900797  
; Patent No. 6645762  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 66457621, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES



Db	2118	AAATATGCGCATCTTCAATGAATATACAGAGATCGAGTGTCATATGACATTTGGGA	2177	Db	238	ACTCCATCTGCTTATGACACAGCATGGGTAGCGAGGTGCCCGCATTTGATGGCTCTGCT	297
Qy	2210	ACAGTCGTAATAAATCAATCTTTTCACTCGAGAAATTAATTCCTGATTAATTAATTAACAG	2269	Qy	494	GCACCCCAATTTCCCAAAAGCGTTCAATGGAATTCGAACAATCAATCCAGATGGCTCG	553
Db	2178	ACAGTCGTTCTCATTAAGTGTCTTTTCACTCGGAGGTTCTTACAGATGAAGTACTCTCC	2237	Db	298	CGCCGCAATTTCCCAAAAGCGTTCAATGGAATTTGAACCAACAGTTAAAGATGGTTCA	357
Qy	2270	CAAGTAGACCTTGGTCCAAATTTCTGATCTTGTCTTTGACTGGAGCACTAATCAAT	2329	Qy	554	TGGGTGAGGAGTGCATTTTTCTCGCGTATGACAGAGTTTAAACACTCTCGCCTGCCTC	613
Db	2238	AAAATTTGATCGCGAATCTAGATTTCTTCAACTCATGCGCTTAAACAGGCGTTTGGTGAAT	2297	Db	358	TGGGGAATTCAGTCCCATTTCTGCTGTCGACCGTCTCTTGGCACTCTTTCTTTGTGT	417
Qy	2330	GACACCAAGACTTACCGCCGAGAGAAACCGTGTGTAATTTGGTTTCCAGCGTACAGTGC	2389	Qy	614	CTCACTCTCAAAATATGAATAAGGCGACATTCAGTGCAGAGAGGGGTTGAGTTTGTG	673
Db	2298	GACACCAAACTTATCAGCAGAGAGAGGTCAAGTGGAGTGGCTTCTGCCATCAATGT	2357	Db	418	CTTGTGCTCTTAAATGGAACGTTGGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA	477
Qy	2390	TACATGAGGGAATCCCGAGTGACAGAGGAAGCTCTAAGTCATGTTTATGGTATC	2449	Qy	674	AGAAACACATGGAAGAAATGAAGCAAGCTGACAATCAAG---GCCAAGTGGATTC	730
Db	2358	TATATGAAGGACCATCTTAAATCTCTGAAGAAGAGCTCTACAAATGTCTTATAGTGC	2417	Db	478	AAGAGCAATCTGGAACTAGTAAGGATGAACCGATCAAGATAGCTTGTGTAACAGACTTT	537
Qy	2450	ATCGACACGCACTGAAGAAATTTGAGTGGAGTTGGCCACCCAGCGAGCAATGCCCA	2509	Qy	731	GAGGTGCTGTTTCTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATCTTCTTAT	790
Db	2418	ATGGAATATGCGCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAAATACGG	2471	Db	538	GAGATCATATTTCTCTCTGTTAAGAGAAGCTCAATCTCTGCGCTCGGACTTCCCTAC	597
Qy	2510	TTGTGCTGAGAGAGTCTGTTTCAACACTGCAAGAGTGAATGCACTGTTTATATGTAC	2569	Qy	791	CACCTCCCTTTCATCTCCCAATCCACCAGCCAGAGAAAAGCTTCAAAAGATTTCC	850
Db	2472	GATATTTACAAAGACTGGTTTGGAACTGCAAGAAATATGCAACTCTTTTATATGCAA	2531	Db	598	GACCTGCTTATATACATCTGTTCAGACTTAACCGCAGGAAAGATTAGCAAACTTTCA	657
Qy	2570	AGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTGACGCAACTCTT	2626	Qy	851	CTCAATGTTCTTCAATAACCATCAGACGGGTTGCTCTACTCTCTCGAGGGTTTGAAGAT	910
Db	2532	GGGGATGGTTTGACACTATACATGATATGGAATTTAAAGAGCATGTCAAAAATGCGCTC	2591	Db	658	AGGAGGAATTTATGCGGTCTCGCTGCGCTTGTGTTATCTTTAGAGGGAATACAAGAT	717
Qy	2627	TTTCGATCTGTCGCGTAGCATA	2648	Qy	911	GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGCGGATCATTTTNAAGCTCC	970
Db	2592	TTCCACCAAGTTGCGCTAGATTA	2613	Db	718	ATAGTTGAATGGGAACGAATTAATGGAAGTTCAAGTCCAGGATGGGTCTTTCTTAAAGCTCA	777
RESULT 6							
US-09-234-393-39							
; Sequence 39, Application US/09234393A							
; Patent No. 6265639							
; GENERAL INFORMATION:							
; APPLICANT: Croteau, Rodney B							
; APPLICANT: Crock, John E							
; APPLICANT: Bohlman, Jorg							
; APPLICANT: Jetter, Reinhard							
; APPLICANT: Steele, Christopher L							
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)							
; TITLE OF INVENTION: AND METHODS OF USE							
; FILE REFERENCE: WSUR113345							
; CURRENT APPLICATION NUMBER: US/09/234,393A							
; CURRENT FILING DATE: 1999-01-20							
; EARLIER APPLICATION NUMBER: 60/072,204							
; EARLIER FILING DATE: 1998-01-22							
; NUMBER OF SEQ ID NOS: 55							
; SOFTWARE: PatentIn Ver. 2.0							
; SEQ ID NO 39							
; LENGTH: 2525							
; TYPE: DNA							
; ORGANISM: Artificial Sequence							
; FEATURE:							
; OTHER INFORMATION: Description of Artificial Sequence:							
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding							
; OTHER INFORMATION: E-alpha-bisabolene synthase							
; FEATURE:							
; NAME/KEY: CDS							
; LOCATION: (1)..(2451)							
US-09-234-393-39							
Query Match 27.0%; Score 731.6; DB 3; Length 2525;							
Best Local Similarity 59.7%; Pred. No. 3.2e-210;							
Matches 1369; Conservative 0; Mismatches 894; Indels 31; Gaps 7;							
Qy	434	AAATCATCTGCATATGATACAGCTTGGGTGGCAGAAATTCGGTCAATTTGACGGCTCTGCT	493	Qy	1511	AAGGATCTTCCTGGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAGTATGCCA	1570

1315 CAAACCTGAGCCAGAGATCAAAATACGCGTGAAGACTTCTTTGGCATGCCAGTGTTCG 1374  
1571 AGATTGAGGCAAGAGTTACATAGACCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG 1630  
1375 AGAGTGAAGCAAGAGATATCTGTCAAGTGTATCGCCAGATATATGACGCATAGCAAAA 1434  
1631 ACTGTGTATAAGATGCTATATGTGAGCAACGAAATAATTTGGAGCTGGCCAAATTTGGAC 1690  
1435 TGGTTTTACAGCTACCTACGTGGAACAAATGAAAGTTTTTGTAGCTGGGAAATTAGAT 1494  
1691 TTCAATATGTTGACGGCTTTACACCAAAAGGAGACTCAACACATTTGTACGCTGGTGAGA 1750  
1495 TTCAACATTATCCAGTCCATCCACCAAGAAGAAATGAAGATGTTACCAAGCTGGTTAGA 1554  
1751 GAATCGGATTCATATGATCTTACATTCACCGCCAGCGGCTGTGGAAATGATTTCTCA 1810  
1555 GATTCGGGGTTGCCATATTTACACCTTCGCTCGGAGAGGCGCTGGAAATTTCTATCTTTA 1614  
1811 GTGCGGTTAGTATGTTTGGAGCCAGAAATTCGCTGTTGTAGAATTCGCTATGCAAGACT 1870  
1615 GTAGCGCGGGACCTATGAACCCAGTATGCCAAATGCGAGTTCTCTTTTACAAAGTG 1674  
1871 TCTTGCTCGCAGTTATTTAGACGATCTTTAGCACCCACCGATCTCTGATGATCTT 1930  
1675 GCATGCTTCAGACTGTTCTGGAGCATATGTATGACACTTATGGAACCTTAGATGATG 1734  
1931 AATTTGTTCTGTAAGCGGTCGGAAGATGGATATCTCTGTCTGGATAGCGTTGGGAT 1990  
1735 AAGCTATTTCACTGAGGCTGTGAGAAGATGGACCTCTCTTTTACAGAAACCTTCCAGAC 1794  
1991 AATCAGTTGAAGTTGCTTCTTAGGCTGTACAAACACAGTGAATTTGGAAAGAT 2050  
1795 TAT---ATGAACATATGTTACCAATCTATATGACATAGTTACAGAGTGGCTTGGAG 1851  
2051 GGAATCAAGGAACAGCGCTGATGTCTGGCTATCTTCGAAAGTATGGAGGGCTTG 2110  
1852 GCAGAGAGGACAGGGCGTGAATTTGTCAGCTTTTTCAGAAAGGATGGAGGATTAT 1911  
2111 CTGCAATCTATACCAAGAGCCGAATGTTGCGCAGCAAGATGATGTGCGGACATTTCAAC 2170  
1912 CTTCTGGGTTATTATGAAGAAGCTGAATGGTTAGTCTGATGATGTGCTACCTTGGAC 1971  
2171 GAATATGTGAATATGCAAGTGTCCATAGCACTTCGGACAGTCTGTAATCAATCAATC 2230  
1972 GAGTACATAAAGAAATGGAATCACATCTATCGGCCAACGTATATCTTCTGTTGAGTGGAGT 2031  
2231 TTTTTCAT---CTGAGAAATTAATCTCTGATTTACATTTTACAGCAAGTAGACCTTCGGTCC 2287  
2032 TTGNTAATGATGAGGCAATCTCTTTCGCAAGAGCATTAGAAGATGATATTCAGGA 2091  
2288 AAATTTCTGC-----ATCTTGTGTCTTTGACTGGACGACTAATCAATGACACCAAGACT 2341  
2092 AGACGTGTTCTCACAGAGCTGAATAGCTCATTTTCCCGCTCGGGATGACACGAAGACA 2151  
2342 TACAGCGCGAGAAACCGTGGTGAATTTGTTTCCAGCGTACAGTCTCATGAGGGAA 2401  
2152 TATAAAGCTGAGAGGCTCGTGGAGAAATGGCGTCCAGCATTTGAATGTTACATGAAAGAC 2211  
2402 AATCCGGAGTGCACAGAGGAAGAGCTTAAGTTCATGTTATGTTATCATCATCGACCA 2461  
2212 CATCTGAATGTAACAGAGGAAGAGCTCTCATCATCATATATAGCATCTGGAGCCGCG 2271  
2462 CTGAAGAAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2521  
2272 GTGAAGGAATCTGACAGAGAGTTTCTGAAGCCCG---ACGACGTCCCATTTCCGCTGCAAG 2328  
2522 AGACTGCTGTTCAACACTGCAAGATGATGACAGCTGTTTATATATGATGACAGATGGCTTT 2581  
2329 AAGATGCTTTTTCGAGGAGACAAAGAGTGCACGATGGTGATATTTCAAGAGATGGAGATTC 2388  
2582 GGTATCTCTGCAAGAGATGAAAGACCATGTACAGCCGAACTCTTTTTCGATCTCTGTGCGG 2641  
2389 GGTGTTTCCAAATTAGAAGTCAAAGATCATATCAAGAGTGTCTCATTTGAACCCCTGCCA 2448

Qy 2642 TAGCA-----TACTGATATATATATATATATTTTCAATTCACCAAAAAA 2691  
Db 2449 CTGTAATCAAAATAGTTGCAATTAATAATTTGAATATCAACTATGTTTTCACAAAAA 2508  
Qy 2692 AAAAAA 2705  
Db 2509 AAAAAA 2522

## RESULT 7

US-09-865-171-39  
; Sequence 39, Application US/09865171  
; Patent No. 6451576  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B  
; APPLICANT: Crock, John E  
; APPLICANT: Bohlmann, Jorg  
; APPLICANT: Steele, Christopher L  
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),  
; FILE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: WSUR117468  
; CURRENT APPLICATION NUMBER: US/09/865,171  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/234,393  
; PRIOR FILING DATE: 1999-01-20  
; PRIOR APPLICATION NUMBER: 60/072,204  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 2525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding  
; OTHER INFORMATION: E-alpha-bisabolene synthase  
; NAME/KEY: CDS  
; LOCATION: (1)..(2451)  
US-09-865-171-39

Query Match 27.0%; Score 731.6; DB 4; Length 2525;  
Best Local Similarity 59.7%; Pred. No. 3.2e-210;  
Matches 1369; Conservative 0; Mismatches 894; Indels 31; Gaps 7;

Qy 434 AATCCATCTGCATATGATACAGCTTGGTGGCAGAAATTCGTCGAATTTGACGGCTCTGGT 493  
Db 238 ACTCCATCTGTTATGACACAGCATGGTAGCGGGTGCCTGCCCATTTGATGGCTCTGCT 297  
Qy 494 GCACCCCAATTTCCCAACCGTTTCAATGGAATTCGAAATCAATCACTGCCAGATGGCTCG 553  
Db 298 CGCCCGCAATTTCCCAACAGATTTGACTGGATTTGAAAAACCAAGTTAAAGATGGTTCA 357  
Qy 554 TGGGTTGAGGAGTGCAATTTTCTGGGCTATGACAGAGTTTAAACACTCTCGCTCGCTC 613  
Db 358 TGGGGAATTCAGTCCCACTTTCTGCTGCCGCGTCTTCTTGGCACTCTTCTTGTGTT 417  
Qy 614 CTCACCTCTCAAAATATGGAATAAGGGCGCATTCAGTGCAGAAAGGGTTGAGTTTGTG 673  
Db 418 CTTGTGCTCTTAATGGAACGTTGGGATCTGCAAGTAGAGCAGGAATTTGAATTCATA 477  
Qy 674 AGAAAAACATGGAAGAAATGAAAGGAGCAAGCTGCAATCACAG---GCCAAGTGGATTC 730  
Db 478 AAGAGCAATCTGGAATCTAGTAAGGATGAAACCGATCAAGATAGCTTTGGTAACAGACTTT 537  
Qy 731 GAGTGTGTTTCTGCAATGTTAGATGAAGCAAAAGCTTGGATTTGGATCTTCCCTAT 790  
Db 538 GAGATCATATATTTCTTCTCTGTTAAGAGAGAGCTCAATCTCTGCGCCTCGGACTTCCCTAC 597  
Qy 791 CACTCCCTTTTCACTCTCCCAATCCCAAAAGCGCCAGAAAAAGCTTTCAAAAGATTTCCC 850  
Db 598 GACTGCTTATATACATCTGTTGCAGACTTAACGGCAGGAGAGATTAGCAAACTTTCA 657



QY	851	CTCAATGTTCTTCATAAACCATCAGACGGGTGCTCTACTCTCTCGAGGGTTTGCAAGAT	910
Db	658	AGGGAGGAAATTTATCGGTTCCGTCCCATTTGTGTATCTTTTAGAGGGAATACAAGAT	717
QY	911	GTGGTGCAGCTGGCAAGAGATCACAAATCTTCAATCAAGAGACGGATCATTTTTAAAGCTCC	970
Db	718	ATAGTTGAATGGGAACGAATATGGAGTTCAAAGTCAGGATGGGTCTTTCTTAAAGCTCA	777
QY	971	CCTGCATCTACTGCTGTGTTCATGACACACTCAAAACAAACGATGCCTCCACTTCTC	1030
Db	778	CCTGCTCTACTGCCCTGCGTTTTTCATGCAACACAGGAGACCGGAAATGGCTTGAATCTCTG	837
QY	1031	AACCTCGTCTCAGCAAAATTTGGCGACTAGTTCCTTGGCCATTAACCCACTTCATCTATTT	1090
Db	838	AACAGTGTGAATCAAGTTTGGAAATTTTGTTCCTGCTGTATCTCTGTGATCTGCTG	897
QY	1091	GAACGCCCTCTGGGCTGTGATACAGTTGAAACGCTTCGGGAATCGATGCTATTTCAAGAAA	1150
Db	898	GAACGCCCTGTGATCGTAGATAATATTTGACGCCCTTGGAAATCTATAGACACTTTGAAAAG	957
QY	1151	GAATCAAGAAATCTCTGGATTACGTTTATAGTACTGGGACGCCGAAGAGCGGTGGGA	1210
Db	958	GAATCAAGGAAGCTCTTGATTATGTTTACAGGCATTTGGAACG---AAAGAGGAATTTGGG	1014
QY	1211	TGGSCAAGATGCAATCCTATTTCTCGATGTCGATGACACTGCCATGGGTCTTAGAATCTCTG	1270
Db	1015	TGGGGCAGACTAAATCCATAGCAGATCTTGACCACTGCTTTGGGATTTGATGCTTT	1074
QY	1271	AGACTTCATGGATACAATGATATCTTCAGATGTTCTGGAGAAATTTACAGACGAGAAGA	1330
Db	1075	CGGCTGCATAGGTACAATGATATCTCCAGCCATTTTTCACAACTTTCAAGATGCCAATGGG	1134
QY	1331	GACTCTCTTTGCTTTGCGGTCAAACGCAAAATTTGGTGTGACCGATATCTTAAACCTTTAT	1390
Db	1135	AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTACCAAGCATGCTGATCTTTAT	1194
QY	1391	AGATGTTTCAAGAATGATGTTTTCCGGGAGAAAAGATAATGGAAGAAGCTAAGACCTTCAC	1450
Db	1195	AGAGCTCCCGAGCTCGCATTTCCCGGAGAAAACATTTCTGATGAAGCTAAAGACTTCGCT	1254
QY	1451	ACAAATCATCTC AAAATGCTCTTGCCAAAACAAACGCAATTTGATAGTGGGCTGTCAAG	1510
Db	1255	ACTAAATATTTGAGAGAAGCTCTTTGAGAAAAGTGAGACTTCCAGTGCATGGAACAAACAA	1314
QY	1511	AAGGATCTTCTCGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAATATGCCA	1570
Db	1315	CAAAACCTGACCAAGAGATCAAAATACGCGCTGAAGACTTCTTGGCATGCCAGTTCGG	1374
QY	1571	AGATTGGAGGCAAGAAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG	1630
Db	1375	AGAGTGGGAAGCAAAAGAGATACGTGTCAAGTGTATCGCCACAGATTTATGCACGCATAGCAAAA	1434
QY	1631	ACTGTGTATAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGGCCAAATGGAC	1690
Db	1435	TGCGTTTACAAGCTACCTCAGTGAAACATGAAAAAGTTTTTAGAGCTGGGAAAATTAGAT	1494
QY	1691	TTCAATATGTGTGCGGCCCTTACACCAAAAGGAGACTCAACATGTTGCTAGCTGGTGAGA	1750
Db	1495	TTCAACATTTCCAGTCCATCCACCAGAGAAATGAAGAATGTATACAGCTGGTTTAGA	1554
QY	1751	GAATCGGAATTCATATGATCTTACATTTACCCCGCAGCGGCTGTGGAAATGATTTCTCA	1810
Db	1555	GATTCGGGGTTGCCACTATTCACCTTCGCTCGGAGAGCGCGCTGGAAATTTCTACTCTTA	1614
QY	1811	GTGGCGGTTAGTATCTTTGAGCCAGAAATTCGCTGCTGTGAAATTGCCATATGCCAAGACT	1870
Db	1615	GTAGCGGGGACCTATGAACCCAGTATGCCAAATGCAGGTTCTCTTTTACAAAAGTG	1674
QY	1871	TCTTCGCTCGCAGTTATTTCTAGACGATCTTTACGACCCACGAGATCTCTGGATGATCTTT	1930
Db	1675	GCATGCTTCGACATGTTCTGACCATATGATGACACTTATGAAACCCTAGATGAATTG	1734

QY	1931	AAATTGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTCTCGATAGCCGTTCCGGAT	1990
DB	1735	AAGCTATTCACTGAGGCTGTGAGAAGATGGGACCTCTCTCTTTACAGAAAACCTTCAGAC	1794
QY	1991	AATCAGTTGAAAGTTTGCCTTCCTAGGCGTGTAACAACACAGTGAATGGATTTGGAAAAGAT	2050
DB	1795	TAT---ATGAAACTATGTTTACCAAAATCTATTATGACATAGTTTCAAGAGTGGCTTTGGAG	1851
QY	2051	GGACTCAAGGAAACAGGCGCGTGATGTCTGGGCTATCTTCGAAAAGATATGGGAGGCGTTG	2110
DB	1852	GCAGAGNAGGAAACAGGGCGTGAATTTGGTCAGCTTTTCAGAAAAGGATGGGAGGATTAT	1911
QY	2111	CTCGATCTGATATACAAAGACCGGAATGGTCCGACAGAAAGTATGTGCGGACATTCAC	2170
DB	1912	CTTCTGGGTTTATTATGAAGAAGCTGAATGGTTAGCTGTGAGTATGTGCTACCTTTGGAC	1971
QY	2171	GNATATGTGAAATGCCAAAGTGTCCATAGCACTTTCGACAGTCCGTACTTAAACTCAATC	2230
DB	1972	GAGTACATAAAGAAATGGAATCACATCTATCGGCCAAGGTATACTTCTGTTGAGTGAGTG	2031
QY	2231	TTTTTTCA---CTGGAGAAATTACTTCCTGATTAATTTTTACAGCAAGTAGACCTTCGGTCC	2287
DB	2032	TTGATAATGATGGCAACTCCTTTCCGAAGGCAATTAGAGAAAGTAGATTATCCAGGA	2091
QY	2288	AAATTTCTGC-----ATCTTGTGTCTTTGACTGGAGCACTAATCAATGACCAAGACT	2341
DB	2092	AGACGTGTTCTCAACAGAGCTGAATAGCCTCAITTCGCCCTTCGGGAGTACACGAGACA	2151
QY	2342	TACCAGCGCCGAGAGAAACCGTGGTGAATTTGGTTTTCAGCGCTACAGTGCTCATGAGGGAA	2401
DB	2152	TATAAAGCTGNAAGGCTCGTGGAGAAATTCGGCTCCAGCATTTGAATGTTACATGAAGAC	2211
QY	2402	AATCCGAGTGCACAGAGGAAAGACTCTAAGTCACTGTTTATGGTATCATCGACACGCA	2461
DB	2212	CATCCTGAATGTACAGAGGAAAGGCTCTCGATCACATCTATAGCATCTTGGAGCCGGCG	2271
QY	2462	CTGAAGGAATTTGAATTCGGAGTTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA	2521
DB	2272	GTGAAGGAACTGACACAGAGAGTTTCTGAAGCCG---ACGAGTCCCATTCCTCGCTGCAAG	2328
QY	2522	AGACTGCTGTTTCAACACTGCAAGAGTGATGACGCTGTTTTATATGTACAGAGATGGCTTT	2581
DB	2329	AAGATGCTTTTCGAGGAGACAAGAGTGACGATGGTGATATTCAAGGATGGAGATGGAATTC	2388
QY	2582	GGTATCTCTGACAAAGAGATGAAGACCATGTTCAGCCGAACCTCTTTTCGATCCTGTGCGG	2641
DB	2389	GGTGTCTTCCAAATTAGAAGTCAAAAGATCATATCAAGAGTGTCATCTGAACCCGCTGCCA	2448
QY	2642	TAGCA-----TACTGTATATTATATAATAATTTTCAATTCATATCCAAATCAAAAAA	2691
DB	2449	CTGTATNATCAAAATAGTTGCCAATATTAATTTGAAATATCAACTATGTTTTACAAAAA	2508
QY	2692	AAAAAAAAAAAAAA	2705
DB	2509	AAAAAAAAAAAAAA	2522

RESIST. T. R.

; Sequence 1, Application US/09234393A  
 ; Patent No. 6265639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney B  
 ; APPLICANT: Crock, John E  
 ; APPLICANT: Bohlman, Jorg  
 ; APPLICANT: Jetter, Reinhard  
 ; APPLICANT: Steele, Christopher L  
 ; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
 ; TITLE OF INVENTION: AND METHODS OF USE  
 ; FILE REFERENCE: WSUR11345  
 ; CURRENT APPLICATION NUMBER: US/09/234, 393A  
 ; CURRENT FILING DATE: 1999-01-20  
 ; EARLIER APPLICATION NUMBER: 60/072,204

; EARLIER FILING DATE: 1998-01-22

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2424

; TYPE: DNA

; ORGANISM: Abies grandis

US-09-234-393-1

Query Match 27.0%; Score 730.2; DB 3; Length 2424;

Best Local Similarity 59.6%; Pred. No. 8.2e-210;

Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;

Qy	434	AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCGGTCAATTGACGGCTCTGGT	493
Db	134	ACTCCATCTGTTATGACACAGCATGGGTAGCGAGGGTCCCGCCATTTGATGGCTCTGCT	193
Qy	494	GCACCCCAATTTCCCAACCGCTTCAATGGATTTCTGAACAATCAACTGCCAGATGGCTCG	553
Db	194	CGCCGGCAATTTCCCAACAGTTGACTGGATTTTGAACCAAGTTAAAGATGGTTCA	253
Qy	554	TGGGGTGAAGTGCATTTTCTGGCGTATGACAGAGTTTTAAACACTCTCGCCTGCCTC	613
Db	254	TGGGGAATTCAGTCCCACTTTCTGCTGTCGACCGCTTCTTGGCACTCTTCTTGTTGTT	313
Qy	614	CTCACTCTCAAAATATGGAATTAAGGGCGACATTTCAAGTGGCAAGAGGGGTTGATTTGTG	673
Db	314	CTTGCTCTCTTAAATGGAAGCTTGGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA	373
Qy	674	AGAAACACATGGAAGAAATGAAGCAAGCAAGCTGACATCAAG--GCCAAGTGGATTC	730
Db	374	AAGAGCAATCTGGAACATGTAAGATGAACCGATCAAGATAGCTTGGTAACAGACTTT	433
Qy	731	GAGGTCGTGTTTCTGCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGATCTTCTTTAT	790
Db	434	GAGATCAATTTCTTCTCTGTTAAGAGAAGTCAATCTCTGGCGCTCGGACTTCCCTAC	493
Qy	791	CACCTCCCTTTCATCTCCCAATCCAAAGCGCCAGAAAAGCTTCAAAAAGATTTCCC	850
Db	494	GACCTGCCCTTATATACATCTGTGCGAGCTAAACGGCAGGAAGATTAGCAAAACTTTCA	553
Qy	851	CTCAATGTTCTTCAATAACCATCAGACGGCTGTCTACTCTCTGGAGGTTTCAAGAT	910
Db	554	AGGAGGAAAATTTATGCGGTTCCGTCGCCATTTGTTATCTTTAGAGGGAATCAAGAT	613
Qy	911	GTGTGGACTGGCAAGAGATCACAAATCTTCAATCAAGACGGATCAATTTTAAAGCTCC	970
Db	614	ATAGTTGAATGGGAACGAATAATGGAAGTTCAAGTCAAGATGGTCTTCTTAAAGCTCA	673
Qy	971	CTGCACTACTGCTTGTGCTTTCATGCACTCAAAACAAACAGATGCTCCACTTTCTC	1030
Db	674	CTGCTTCTACTGCTGCGTTTTCATGCAACAGAGACGCGAAATGCTTGAATTTCTTG	733
Qy	1031	AACCTGCTCTCAGCAATTTGGCGACTACCTTCTTCCATACCCACTTGAATCTATTT	1090
Db	734	AACAGTGTGATGATCAAGTTTGGAAAATTTGTTCCTCCCTGTTATCTCTGTGGATCTGCTG	793
Qy	1091	GAAACCTCTGGGCTGTGATACAGTTGAAACGCTTGGGAATCGATCGCTATTTCAAGAAA	1150
Db	794	GAAACCTCTGTGATGATGAATATTTGTAAGCTTGAATCTATAGACACTTTGAAGAG	853
Qy	1151	GAAATCAAGAAATCTCTGGATTAAGTTATAGTTAGTTAGTTAGTTAGTTAGTTAGTTAG	1210
Db	854	GAAATCAAGGAAGCTCTTGATTAATGTTTACAGGCAATGGAAAG--AAAGAGGAATGGG	910
Qy	1211	TGGGCAAGATGCAATCTTCTCTGATGTCGATGATGATGATGATGATGATGATGATGATG	1270
Db	911	TGGGCAAGATGCAATCTTCTCTGATGTCGATGATGATGATGATGATGATGATGATGATG	970
Qy	1271	AGACTTCATGATACAAATGATTTCTTCAGATGTTCTGAGAAATTTTCAGAGACGAGAAAGGA	1330
Db	971	CGGCTGCAATAGGTACAAATGATTTCTCCAGCCATTTTTCAGAAATTTTCAGAAATGGG	1030

Qy	1331	GACTTCTTTTGGCTTTCGGGCTCAACCGCAATTTGGTGTGACCGATAATCTTTAACCTTTAT	1390
Db	1031	AAATTTCAATTTGCTCGACCGCTCAATTTCAACAAAGATGTAGCAAGCATGCTGTAATCTTTAT	1090
Qy	1391	AGATGTTTCAAGATGATGTTTTCGGGAGAAAAGATAATGAAGAAAGCTTAAGACCTTCACT	1450
Db	1091	AGAGCTTCCAGCTCGCATTTTCCGGAGAAAACATTTCTTGATGAAGCTTAAAGCTTCGCT	1150
Qy	1451	ACAAATCATCTCCAAATGCTTTTGCACAAAACAAACGATTTGATTAAGTGGGCTGTCAAG	1510
Db	1151	ACTAAATATTTTGAGAGAAGCTCTTGAGAAAAGTGAGACTTCCAGTGATGTAACAAACAA	1210
Qy	1511	AAGGATCTTCTCGAGAGGTGGAGTATGCTATAAGTATCGTGGCATAGAGTATGCCA	1570
Db	1211	CAAAACCTGAGCCAAAGAGATCAAAATCGCGCTGAAGACTTCTTGGCATGCCAGTGTTCG	1270
Qy	1571	AGATTGGAGGCAAGAAATGTACATAGAGCAATTTTGGATCAAAATGATGTCTGGCTGGGAAG	1630
Db	1271	AGAGTGGAGCAAGAGATGATGTCAGTGTATCGCCAGATTTATGCAGCATAGCAAAA	1330
Qy	1631	ACTGTGTATAGATGCTATATGTGAGCAACGAAAATATTTTGGAGCTGGCCAAATTTGGAC	1690
Db	1331	TGCGTTTACAAAGCTTACCTACGTGAAACATGAAAAGTTTTTAGAGCTGGGAAAATTAGAT	1390
Qy	1691	TTCAATATGTGCAAGGCTTACACCAAAAGAGACTCAACACATTTGTCAGCTGCTGGAGA	1750
Db	1391	TTCAACATTTACCAAGCTTCCACCAAGAGAATAAGAAATGTTACAGCTGTTTAGA	1450
Qy	1751	GAATCGGGATTTCAATGATCTTCAATTCACCCCGCAGCGCTGTGGAATGATTTCTCA	1810
Db	1451	GAATCGGGATTTGCACTATTTCACTTCCCTCGCTGGGAGAGCGCTGGAAATCTACTCTTA	1510
Qy	1811	GTGCGGTTAGTATGTTTGGAGCCAGAAATCGCTGCTTGTAGAAATTTGCTATGCAAGACT	1870
Db	1511	GTAGCGCGGGAGCTATGAACCCAGTATGCCAAATGCGAGTTCTCTTTTACAAAAGTG	1570
Qy	1871	TCTTGCCTCGAGCTTATCTTAGAGCATTTTACGACACCCAGCTCTCTGATGATGCTT	1930
Db	1571	GCATGCTTGCAGACTGTTTCTGGACGATATGTAACACTTATGGAACCTTAGATGAAATG	1630
Qy	1931	AAATTTCTCTGAAGCGGTCCGAAGATGGATATCTCTGTGCTGGATAGCGTTCCGGAT	1990
Db	1631	AGCTTATCTAGGCTGTGAGAAATGGAGCTCTCTTTTACAGAAAACCTTCCAGAC	1690
Qy	1991	AATCAGTTGAAAGTTGCTTCTTAGGGCTGTACAAACACAGTGAATGGAATTTGSAAGAT	2050
Db	1691	TAT---ATGAAACTATGTTACAAATCTATATGACATAGTTTACAGAGTGGCTTGGAG	1747
Qy	2051	GGACTCAAGAACAGCGCGGTGATGCTGGGCTATCTTGGAAAAGTATGGAGGGCTTG	2110
Db	1748	GCAGAGAAGAACAGGGGCGTGAATTTGCTCAGCTTTTTCAGAAAAGGATGGAGGATTTAT	1807
Qy	2111	CTGCACTGCTATACCAAGAAAGCGGAATGTTGGCAGCAAGATATGTGCCACATTTCAAC	2170
Db	1808	CTTCTGGGTTATTTAGAAAGCTGAATGGTTAGTCTGCTGATGATGCTTACCTTGGAC	1867
Qy	2171	GAATATGTGGAATAATGCCAAAGTGTCCATAGCACTTCCGACAGTCTATAAATCTCAATC	2230
Db	1868	GAGTACATAAAGAAATGGAATCACAATCTATCGGCCAACGATATCTTCTGTTGAGTGGAGT	1927
Qy	2231	TTTTTCA---CTGAGAAATTTACTTCTGATTTACATTTTACGAAAGTAGAGCTTCCGTCC	2287
Db	1928	TTGATAATGATGGCAACTCTTTTTCGCAAGAGGCAATTAGAGAAAGTAGATTTATCCAGGA	1987
Qy	2288	AAATTTCTGCG-----ATCTTGTGCTTTTGAATGAGCTTAATCAATGACCAAGACT	2341
Db	1988	AGAGCTGTTCTCAGAGCTGAATAGCTCATTTTCCCGCTTGGCGGATGACACGAGACA	2047
Qy	2342	TACCAAGCCGAGAGAAACCTGTTGTAATGTTTTCAGCGGTACAGTGTACATGAGGAA	2401
Db	2048	TATAAAGCTGAGAGGCTCGTGGAGAAATGGCGTCCAGCATTTGAATTTTATCATGAAGAC	2107
Qy	2402	AATCCGAGTGCAACAGAGGAAGAGCTCTTAAGTCTTAAGTCTATGTTATGTTATCATCGAACGCA	2461



```
Db 1451 GATTCGGGGTGCACATTCACCTTCGCTCGGAGAGCGCTGGAAATTCATCTCTTA 1510
Qy 1811 GTGCGGTTAGTATGTTTGAGCCAGAAATTCGCTCTTGTAGAAATGCTATGCCAGACT 1870
Db 1511 GTAGCGCGGGGACCTATGAAACCCAGATGATGCCAAATGCAGGTTCTCTTTACAAAAGTG 1570
Qy 1871 TCTTGCCCTCGCAGTATTCATAGAGATCTTTACAGACCCACCGATCTCTGGATGATCTT 1930
Db 1571 GCATGCTTGCAGACTGTTCTCGACCATATGATGACACTTATGGAACCCCTAGATGAATG 1630
Qy 1931 AAATGTTCTCTGAAGCGGTCGGAAGATGGATATCTCTGTCTGGATAGCTTGGGAT 1990
Db 1631 AAGCTATTCATGAGGCTGTGAGAAGATGGACCTCTCTTTACAGAAACCTTCCAGAC 1690
Qy 1991 AATCAGTTGAAGTTTCTCTAGGCTGTGACAAACACAGTGAATGGAATTTGAAAAAGAT 2050
Db 1691 TAT---ATGAAACTATGTTACCAATCTATTATGACATAGTTACAGAGGTGGCTGGAG 1747
Qy 2051 GGACTCAAGAACAGAGCGGTGATGCTCGGCTATCTTCGAAAAGTATGGAGGGCTTG 2110
Db 1748 GCAGAGAAAGAACAGGGCGGTGAATGCTGAGCTTTTTCAGAAAGGATGGAGGATAT 1807
Qy 2111 CTCGATCGTATACCAAGAGCGGAATGTCGCGACGAAAGTATGTCGACATTCAC 2170
Db 1808 CTTCTGGGTTATTATGAAGAGCTGAATGGTTAGTCTGAGTATGTCCTACCTTGGAC 1867
Qy 2171 GAATATGTGGAAGTCCAAAGTGTCCATAGCACTTCGCGACGTCGTAACCTCAATC 2230
Db 1868 GAGTACATAAAGAAATGGAATCACAATCTATCGGCCAACGTATCTCTGTTGAGTGGAGTG 1927
Qy 2231 TTTTTC---CTGAGAAATTAATCTCTGATTAATTTTACAGCAAGTAGACCTTCGGTCC 2287
Db 1928 TTGATATGATGGAGCACTCTTTTCGCAAGAGCAATTAGAGAAAGTAGATATTCACGGA 1987
Qy 2288 AAATTTCTGC-----ATCTTGTCTCTTCAGTGGAGCACTTAATCAATGACACCAAGACT 2341
Db 1988 AGAGTGTTCTCAGAGCTGAATAGCTCATTTTCCCGCTGGCGGATGACAGACGA 2047
Qy 2342 TACCAGCCGAGAGAAACCGTGTGTAATGTTTTCAGCGTACAGTCTACATGAGGGAA 2401
Db 2048 TATAAAGCTGAGAAGGCTCGTGGAGAAATGCGCTCCAGCAATTGAATGTTACATGAAGAC 2107
Qy 2402 AATCCGAGTGCAAGAGGAAAGCTCTAAGTCAATGTTATGATATCATCGACACGCA 2461
Db 2108 CATCTCTGAATGTAACAGAGGAGGCTCTCGATCACAATCTATAGCAATCTCGAGCCGCG 2167
Qy 2462 CTGAAGAAATGTAATGGAGTGGCCCAACCCAGGAGCAATGCCCATTTGTGTGAGA 2521
Db 2168 GTGAAGAACTGACAGAGAGTTTCTGAAGCCCG---ACGACGTCCCATTTGCGCTGCAAG 2224
Qy 2522 AGACTGCTGTTCAACACTGCAAGAGTGAATGAGCTGTTTATATGTAACAGAGATGGCTTT 2581
Db 2225 AAGATGCTTTTCGAGGAGACAGAGTGAAGTGAATTCAGGATGGAGTGAATTC 2284
Qy 2582 GGTATCTCTGACAAAGAGATGAAGACCATGTACGCGAACTCTTTTCGATCTCTGTGGCG 2641
Db 2285 GGTGTTTTCCAAATTTAGAAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCCA 2344
Qy 2642 TAGCA-----TACTGATATTATATAATTAATTCATATTCAATCCAAAAAAA 2688
Db 2345 CTGTAATCAAAATAGTTGCAATAATAATGAAATAATGTCAACTATGTTTTCACAAAAAA 2404
Qy 2689 AAAAAAAAAAAAAAAA 2705
Db 2405 AAAAAAAAAAAAAAAA 2421
```

## RESULT 10

US-09-865-171-1

; Sequence 1, Application US/09865171

; Patent No. 6451576

; GENERAL INFORMATION:

```
; APPLICANT: Croteau, Rodney B
; APPLICANT: Crook, John E
; APPLICANT: Bohlmann, Jorg
; APPLICANT: Steele, Christopher L
; TITLE OF INVENTION: SESQUIITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),
; TITLE OF INVENTION: AND METHODS OF USE
; FILE REFERENCE: WSUR117468
; CURRENT APPLICATION NUMBER: US/09/865,171
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 09/234,393
; PRIOR FILING DATE: 1999-01-20
; PRIOR APPLICATION NUMBER: 60/072,204
; PRIOR FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Abies grandis
; US-09-865-171-1
```

```
Query Match 27.0%; Score 730.2; DB 4; Length 2424;
Best Local Similarity 59.6%; Pred. No. 8.2e-210;
Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;
```

```
Qy 434 AATCCATCTGCATATGATACAGCTTGGTGGCAAGAAATTCGGTCAATGAGCGCTCTGGT 493
Db 134 ACTCCATCTGTTATGACACAGATGGGTAGCGAGGGTGC CGCATTGATGGCTCTGCT 193
Qy 494 GCACCCCAATTTCCCAACCGCTTCAATGATTTCTGAACAATCAACTGCCAGATGGCTCG 553
Db 194 CGCCCGCAATTTCCCAACAGATGACTGGATTTTGAANAACCAATGTAAGATGGTTCA 253
Qy 554 TGGGGTGAGGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCCTC 613
Db 254 TGGGGAATTCAGTCCCACTTTCTGCTGTCGACGCTCTCTTGCCACTCTCTTCTTGTT 313
Qy 614 CTCACTCTCAAAATATGGAATAAGGGGCAATTCAGTGCAGAAAGGGTGTGATTTGTG 673
Db 314 CTTGTGCTCTTAAATGGAACGTTGGGGATCTCAAGTAGAGCAGGGAATTTGAATTCATA 373
Qy 674 AGAAACACATGAAGAAATGAAGGACGAAGCTGACAAATCACAG---GCCAAGTGGATTC 730
Db 374 AAGAGCAATCTGGAACATAGTAAAGGATGAAACCGATCAAGATAGCTTGGTAAACAGACTTT 433
Qy 731 GAGTCTGTTTCTCTCAATGTTTAGATGAACAAAAGCTTGGGATTTGGATCTTCTCTTAT 790
Db 434 GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC 493
Qy 791 CACTCTCTTTTCATCTCCCAAAATCCACAAAAGCGCCAGAAAAGCTTCAAAAGATTTCCC 850
Db 494 GACCTGCTTTATATACATCTGTTGCAGACTTAAACGGCAGGAAAGATTAGCAAAACTTTCA 553
Qy 851 CTCAATGTTCTTCATACCAATCAGAGCTGCTCTACTCTCTGAGGGGTTTGCAGAGAT 910
Db 554 AGGAGAAATTTATGCGGTTCGCTGCGCAATGTGTATTCTTTAGAGGGAATACAAGAT 613
Qy 911 GTGGTGAGCTGGCAAGAGATCAAAAATCTTCAATCAAGAGACGAGATCAATTTTAAAGCTCC 970
Db 614 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTAAAGCTCA 673
Qy 971 CTTGCAATCTACTGCTGTGTGTTCTTCATGCACTCACTCAAAAACAAACGATGCTCCACTTCTC 1030
Db 674 CTTGCTTCTACTGCTGCTGTTTTCATGCACTCAAGGAGACGCAAAATGCTTTGAATTTCTTG 733
Qy 1031 AACTTGTGCTCAGCAAAATTTGGGCACTAGCTTCTTCCCTTGGCAATACCCACTTGATCTATTT 1090
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGCTATCTCTGTGGATCTGCTG 793
Qy 1091 GAACGCTCTGGGCTGTGCGATACAGTTGAACGCTTGGGAATCCGATCCGCTATTTTCAAGAAA 1150
Db 794 GAACGCTTGTGATCGTAGATAATATTGTAAGCTTGGAAATCTATAGACACTTTTGAAGAG 853
```



Db 134 ACTCAATCTGCTTATGACACAGCATGGGTAGCGAGGGTGCCCGCAATTGATGGCTCTGCT 193  
Qy 494 GCACCCCAATTTCCGCAACGCTTCAATGGAATCTGAAACAATCAACTGCCAGATGGCTCG 553  
Db 194 CGCCCGCAATTTCCGCAACAGATTTGACTGGATTTTGAAGAACCAAGTTAAAGATGGTTCA 253  
Qy 554 TGGGGTGAGGAGTGCAATTTTCTGCGGTATGACAGAGTTTTTAAACACTCTCGCCTGCTC 613  
Db 254 TGGGGAATTCAGTCCCACTTCTGCTGTCCGACGCTTCTTGGCACTCTTCTTGTTGTT 313  
Qy 614 CTCACTCTCAAAATPATGGAATPAAGGGCGACATTCAGGTGCGAGAAAGGGTTGAGTTTGTG 673  
Db 314 CTTGTGCTCTTAATAATGGAACGTTTGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA 373  
Qy 674 AGAAACACATGGAAGAAATGAAGGACGAGCTGACAAATCAAG- - -GCCAAGTGGATTC 730  
Db 374 AAGAGCAATCTGGAACATAGTAAAGGATGAAACCGATCAAGATAGCTTTGGTAAACAGACTTT 433  
Qy 731 GAGGTGCTGTTCTGCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGATCTTCCCTTAT 790  
Db 434 GAGATCAATTTCTCTCTGTTAAGAGAGCTCAATCTCTGCGCCTCGGACTTCCCTAC 493  
Qy 791 CACCTCCCTTTCTCTCCCAATCCAGCAAAAGCGCCAGAAAAGCTTCAAAAAGATCCC 850  
Db 494 GACCTGCTTATATACATCTGTTGCAGACTAAACGGCAGGAAAGATTAGCAAACTTTCA 553  
Qy 851 CTCATGTTCTTCATACCACTCAGACGGCGTTGCTCTACTCTCTGAGGGTTTGCAAGAT 910  
Db 554 AGGAGGAAATTTATGCGGTTCCGTCGCCATTTGTTGATTTCTTTAGAGGGAATACAAGAT 613  
Qy 911 GTGTGACTGGCAAGAGATCACAAATCTTCATCAAGAGACGATCATTTTAAAGTCC 970  
Db 614 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAAGTCAGGATGGGTCTTTCTTAAGCTCA 673  
Qy 971 CCTGCATCTACTGTTGTGCTTCAATGCACACTCAAAACAAACGATGCCCTCCACTTCTC 1030  
Db 674 CTTGCTTCTACTGCTGGTTTTCATGCACACAGGACGCGGAATGCTTGAATTTCTG 733  
Qy 1031 AACTTGTGCTCAGCAATTTGGCGACTAGCTTCTTGGCAATTAACCACTTGATCTAATTT 1090  
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGATCTCTGTGGATCTGCTG 793  
Qy 1091 GAAGCCTCTGGGCTGTGATACAGTTGAACGCTTGGGAATCGATCCCTATTATCAAGAA 1150  
Db 794 GAAGCCTGTGTGCTGTAGATAATATTGTACGCTTGGAAATCTATAGACACTTTGAAAG 853  
Qy 1151 GAAATCAAGAAATCTCTGGATTTACGTTTATAGTACTTGGGACGCGGAAAGCGGTGGGA 1210  
Db 854 GAAATCAAGGAGCTCTTGATTTATGTTTACAGGCAATTGGACG- - -AAGAGGNAATTGGG 910  
Qy 1211 TGGGCAAGATGCAATCTTCTGATGTGATGACACTGCCATGGGTCTTAGAATCCTG 1270  
Db 911 TGGGCGAGACTAAATCCCATAGCAGACTTGTGAGACCCTGCTTTGGGATTTGATTCGTT 970  
Qy 1271 AGACTTCATGATACATGATCTTTCAGATGTTCTGAGAAATTTCAAGACGAGAAAGGA 1330  
Db 971 CGGCTGATAGTACAAATGATCTCCAGCCTTTTGGCAACTTTTCAAAAGATGCCAATGGG 1030  
Qy 1331 GACTTCTTTGCTTTCGCGTCAACGCAATTTGGTGTGACCGGATATCTTAACTTTAT 1390  
Db 1031 AATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGAACGATGCTGAATCTTTAT 1090  
Qy 1391 AGATGTTCAAGATGATGTTTTCGGGAGAAAGATAATGGAAGAGCTAAGACCTTCACT 1450  
Db 1091 AGAGCTTCCAGCTCGCAATTTCCGAGAGAAAACATTTCTGATGAAGCTAAAGCTTCGCT 1150  
Qy 1451 ACAAATCATCTCCAAATGCTCTTGCACAAAACCAACGATTTGATAGTGGCTGTCAAG 1510  
Db 1151 ACTAAATATTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAAACAA 1210  
Qy 1511 AAGGATCTTCTGAGAGAGTGAGATGCTATTAAGTATCCGTGGCATAGAGTATGCCA 1570  
Db 1211 CAAAACCTGAGCCAGAGATCAAAATACGCGCTGAGAGACTTCTTGGCATGCGAGTTCCG 1270

Qy 1571 AGATTGGAGCAAGAGTTTACATAGACCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG 1630  
Db 1271 AGAGTGGAAAGCAAGAGATATCTGTCAAGTGTATCGCCAGATTATGACCGCATAGCAAAA 1330  
Qy 1631 ACTGTGTATAAGATGCTATATGTGAGCAACGAAATAATTTGGAGCTGGCCAAATTTGAC 1690  
Db 1331 TGGTTTCAAGCTACCTACGTGAAACAAATGAAAGTTTTTTAGAGCTGGGAAATTTAGAT 1390  
Qy 1691 TTCAATATGTTGAGGCTTTACACAAAAGGAGACTCAACACATTTGTGAGCTGGTGAGA 1750  
Db 1391 TTCAACATTTATCCAGTCCATCCCAAGAAAGAAATGAAGATGTTTACAGCTGTTTGA 1450  
Qy 1751 GAATCGGATTTCAATGATCTTACATCCCGCCAGCGGCTGTGGAAATGTAATTTCTCA 1810  
Db 1451 GATTTCGGGGTTGCCACTATTTTACCCTTCGGGAGAGCGCTGGAAATTTCTACTTCTTA 1510  
Qy 1811 GTGGCGTTAGTATGTTTGGAGCAGAAATTCGCTGCTGTAGAAATTTGCTTATGCCAAGACT 1870  
Db 1511 GTAGCGCGGGAGCTTATGAAACCCAGTATGCCAATGCAAGTTTCTCTTTTACAAAAGTG 1570  
Qy 1871 TCTTGCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCACCGATCTCTGGATGATCTT 1930  
Db 1571 GCATGCTTGACAGCTGTTCTGGACGATATGTATGACACTTTATGGAACCTTAGATGAATTG 1630  
Qy 1931 AAATGTTCTCTGAAGCGGTCCGAAAGATGGGATATCTCTGTGTGGATAGCGTTTCGGAT 1990  
Db 1631 AAGCTATTCTCAGGCTGTGAGAAGATGGGACCTCTCTTTTACAGAAAACCTTCCAGAC 1690  
Qy 1991 AATCAGTTGAAATTTGCTTCTTAGGCTGTACACACAGTGAATGGAATTTGGAAGAGAT 2050  
Db 1691 TAT- - -ATGAAACTATGTTTACCAAATCTATTATGACATAGTTTACAGAGTGGCTTGGAG 1747  
Qy 2051 GGAATCAAGGAAACAGCGCTGATGTGCTGGGCTATCTTCGAAAAGTATGGGAGGCTTG 2110  
Db 1748 GCAGAGAGGAAACAGGGCGTGAATTTGCTCAGCTTTTTCAGAAAGGATGGAGGATTTAT 1807  
Qy 2111 CTGCACTGATATACAAAGAGCGGATGTTGCGGAGCAAAAGTATGTGCGGACATTTCAAC 2170  
Db 1808 CTTCTGGGTTTATTTATGAAGAGCTGAATGGTTAGTGTGCTGAGTATGTGCTTACCTTGAC 1867  
Qy 2171 GAATATGTGAAATCCAAAGTGTCCATAGACTTTCGAGACGCTGCTACTAAACTCAATC 2230  
Db 1868 GAGTACATAAAGAAATGGAATCAATCTATCGGCCAAAGTATATCTTCTGTTGAGTGGAGTG 1927  
Qy 2231 TTTTTC- - -CTGGAGAAATTTACTTCTCTGATTATCAATTTTACAGCAAGTAGACCTTCGGTCC 2287  
Db 1928 TTGATTAATGGATGGGCAACTCTCTTTCGCAAGAGGCATTAGAGAAAGTAGATTATCCAGGA 1987  
Qy 2288 AAATTTCTGC- - -ATCTGTGTCTTTGACTGGAGCACTAATCAATGACACCAAGACT 2341  
Db 1988 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCCGCTGGCGGATGACACGAGACA 2047  
Qy 2342 TACCGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTCTACATGAGGGA 2401  
Db 2048 TATAAGCTGAGAGGCTCGTGAGAAATTTGGCGTCCAGCAATTTGAATTTTATCATGAAAGAC 2107  
Qy 2402 AATCCGAGTGACAGAGGAAGAGCTTAAGTCAATGTTTATGTTATCATCGACCAACGCA 2461  
Db 2108 CATCTCGAATGTACAGGAGAGGCTCTCGATCAATCTATAGCATTTCTGGAGCCGCG 2167  
Qy 2462 CTGAAGAAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2521  
Db 2168 GTGAAGAACTGACAAAGAGAGTTTCTGAAGCCG- - -ACGACGTCCCATTTCCGCTGCAAG 2224  
Qy 2522 AGACTGCTGTTCAACTGCAAGAGTGTGACGCTGTTTATATGTATACAGAGATGGCTTT 2581  
Db 2225 AAGATGCTTTTCGAGGAGCAAGAGTGCAGATGGTGAATTTCAAGGATGGAGATGGAATC 2284  
Qy 2582 GGTATCTCTGACAAAGAGATGAAAGACCATGTGCGCGAACTCTTTTCGATCTCTGTGCGG 2641  
Db 2285 GGTGTTTCCAAATTTAGAAATCAAGATCATATCAAGAGATGTTCTCATTTGAACCGCTGCCA 2344







Db	911	TGGGGCAGACTAAATCCCATAGCAGATCTTTGAGACCACTGCTTTGGGATTTGCATTTGCTT	970
Qy	1271	AGACTTCATGGATACAATATGATCTTTCAGATGTTTCTGGAGAAATTTACAGACGAGAAAGA	1330
Db	971	CGGCTGCATAGGTACAAATGATCTCCAGCCATTTTTTGACAACTTCAAGATGCCAATGGG	1030
Qy	1331	GACTTCCTTTTGGCTTTGCCGGTCAAAACGCAAAATTTGGTGTGACCGATTAATCTTAACTTTAT	1390
Db	1031	AAATTTCAATTTTGCCTCGAACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1090
Qy	1391	AGATGTTTCAAGATGATGTTTTCGGGAGAAAGATAATGGAAGAACTAAGACCTTCACT	1450
Db	1091	AGAGCTTCCAGCTCGCATTTCCCGNGAAGAAACATTTCTTGAAGAAGCTAAAGCTTCGCT	1150
Qy	1451	ACAAATCATCTCCAAAATGCTCTGTGCCAAAAACAACGCAATTTGATTAAGTGGGCTGTCAAG	1510
Db	1151	ACTAAATATTTGAGAGAAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAAACAACAAA	1210
Qy	1511	AAGGATCTTCTGGAGAGGTGGAGTAGTATAAAGTATCCGGTGATAGAGTATAGCAATGCCA	1570
Db	1211	CAAACTGTAGCCCAAGAGATCAAAATCGCCCTGAAGACTTCTTTGGCATGCCAGTGTCCG	1270
Qy	1571	AGATTGGAGCGCAAGAGTTACATAGAGCAATTTGGATCAAATGATCTGGCTGGGGAG	1630
Db	1271	AGAGTGAAGCAAGAGATACTGTCAAGTGTATCGCCAGATTTATGCAACGATAGCAAAA	1330
Qy	1631	ACTGTGTAATAGATCTATATGTAGCAACGAAATAATTTGGAGCTGGCCAAATTTGGAC	1690
Db	1331	TGCGTTTACAAGCTTACCCTACGTGACAAATGAAAGTTTTTTAGAGCTGGGAAATTTAGAT	1390
Qy	1691	TTCAATATGGTGCAGGCTTACACAAAAGAGAGACTCAACAATGTGTCACTGGTGGGAGA	1750
Db	1391	TTCAACATTTCCAGTCCATCCACCAAGAGAAATGAAGAATGTTACCAGCTGTTTTAGA	1450
Qy	1751	GAATCGGAGTTCAATGATCTTATAGACGATCTTTTACGACACCCACGGATCTCTGGATGATCTT	1810
Db	1451	GATTCGGGGTTGCCACTATTCACCTTCGCTCGGAGAGCGCTGGAAATTTCTACTTCTTA	1510
Qy	1811	GTGCGGTTAGTATGTTTGAAGCAGAAATTCGCTCTGTAGAAATGGCTATGCCAAGACT	1870
Db	1511	GTAGCGGGGACCTATGAACCCAGTAGTGCCAATGCAAGTTCTCTTTACAAAGTG	1570
Qy	1871	TCCTTCCCTCGAGTTATCTTAGACGATCTTTTACGACACCCACGGATCTCTGGATGATCTT	1930
Db	1571	GCATGCTTGACAGCTGTTCTGGACGATATGATGACACTTATGACACCCCTAGATGAATG	1630
Qy	1931	AAATTTGTTCTGAGCGGTCCGAAGATGGGATATCTCTGTCTGGATACGTTTCGGAT	1990
Db	1631	AAGCTATTTACTGAGGCTGTGAGAGATGGGACCTCTCTTTACAGAAAACCTTCCAGAC	1690
Qy	1991	AATCAGTTTGAAGTTTGCTTCTTAGGCGTGTACAAACACAGTGAATGGATTTGGAAAAGAT	2050
Db	1691	TAT---ATGNAACTATGTTACCAAACTATTTATGACATAGTTTACAGAGGTGGCTTGGAG	1747
Qy	2051	GGACTCAAGGAAACGAAGCCGTGATGCTGGGCTATCTCTGAAAAGATGTGGAGGGGCTTG	2110
Db	1748	GCAGAGAGGAAACAGGGCGGTGAATTTGTGTACGCTTTTTTCGAAAAGGATGGGAGGATAT	1807
Qy	2111	CTCGCATCGTATACCAAGAAAGCCGAATGGTTCGCGACGAAAGTATGTGCCGACATTCACAC	2170
Db	1808	CTTCTGGGTTTATGAAGAGAGCTGAATGGTTAGCTGTGAGTATGTGCCTACCTTGGAC	1867
Qy	2171	GAATATGTGGAAAATGCCAAAGTGTCCATAGCACTTGGGACAGTCTGTAATAACTCAATC	2230
Db	1868	GAGTACATAAAGAAATGGAATCACATCTATCGGCCAAACGTATCTCTGTTGAGTGGAGTG	1927
Qy	2231	TTTTTCA---CTGGAGAAATTACTCTCGATTTACATTTTACAGCAAGTAGACCTTCGGTCC	2287
Db	1928	TTGATAATGGATTGGCAACTCTCTTCGACAGGCAATTAGAAGAAGTAGATTATCCAGGA	1987
Qy	2288	AAATTTCTGC-----ATCTGTGTGCTTTTGACTGGACGACTTAATCAATGACCAACAGACT	2341
Db	1988	AGACGTGTTCTCACAGAGCTGAATAGCTCTATTTCCCGCTTGGGGGATGACACGAGACA	2047

## RESULT 14

```

US-09-903-012B-45
; Sequence 45, Application US/09903012B
; Patent No. 6569656
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 65696561, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012B
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(2347)
; OTHER INFORMATION: E-alpha-bisabolene synthase
US-09-903-012B-45

```

	Query Match	27.0%	Score	730.2	DB 4	Length	2424
	Best Local Similarity	59.6%	Prod. No.	8.2e-210			
	Matches 1370	Conservative	0	Mismatches	893	Indels	34
						Gaps	7
QY	434	AATCCATCTCGATATGATACAGCTTGGGTGGCGAAGAAATTCGGTCAATTGACGGCTCTGGT	493				
DB	134	ACTCCATCTGCTTATGACACAGCATGGGTACGAGGGTGGCCGCCATTTGATGGGCTCTGCT	193				
QY	494	GCACGCCCAATTTCCCCAAACGCTTTCAATGGATTCTGAACAATCAACTGCCAGATGGGCTCG	553				
DB	194	CGCCCGCAATTTCCCAACACAGTTTGACTGGATTTTGAANAACCAAGTTTAAAGAAGTGGTTTCA	253				

QY 554 TGGGTGAGGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCCTC 613  
Db |||||  
QY 254 TGGGGAATTCAGTCCCACTTCTGCTGTCGACCGTCTCTTGCCACTCTTCTTGTT 313  
Db |||||  
QY 614 CTCACCTCTCAAAATATGGAATAAGGCGCAGCATTTCAAGTGCAGAAAGGGTTGAGTTGTG 673  
Db |||||  
QY 314 CTTGTGCTCTTAAATGGAACGTTGGGATCTGCAAGTAGAGCAGGGAATTGAATTCATA 373  
Db |||||  
QY 674 AGAAACACATGGAAGAAATGAAGAGCAAGAGCTGACAACTACAG- - -GCCAAGTGAATTC 730  
Db |||||  
QY 374 AAGAGCAATCTGGAATCTAGTAAGAGATGAAACCGATCAAGATAGCTTGGTAACAGACTTT 433  
Db |||||  
QY 731 GAGGTGCTGTTCTCCTGCAATGTTAGTAGAGCAAAAGCTTGGATGTTGATCTTCTTAT 790  
Db |||||  
QY 434 GAGATCATATTTCTCTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC 493  
Db |||||  
QY 791 CACCTCCCTTTTCATCTCCCAATCCACAAAGCGCCAGAAAAGCTTCAAAAGATTTCC 850  
Db |||||  
QY 494 GACCTGCTTATATACATCTGTTGCAGACTAAACGGCGAGAAAGATTAGCAAACTTCA 553  
Db |||||  
QY 851 CTCAAATGTTCTTCAATACCATCAGACGGGTTGCTCTACTCTCTGAGGGTTTGAAGAT 910  
Db |||||  
QY 554 AGGAGGAAATTTATGGGTTCCGTCGCCATTTGTTATTTCTTTAGAGGGAATACAAGAT 613  
Db |||||  
QY 911 GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGACCGATCATTTTAAAGTCC 970  
Db |||||  
QY 614 ATAGTTGAAATGGGAACGAATATGGAAGTTTCAAGTCAGGATGGGCTCTTTCTTAAGCTCA 673  
Db |||||  
QY 971 CTTGCACTACTCTGCTGTTCTTATGCACACTCAAAACAAAGATGCTCCCACTTTCTC 1030  
Db |||||  
QY 674 CTTGCTTCTACTGCTGCTGTTTCTATGCAACAGGAGACGGGAATGCTTGAATTTCTTG 733  
Db |||||  
QY 1031 AACTTCGTCTCAGCAAAATTTGGCACTACGTTCTTGGCCATTTACCCACTTGATCTATTT 1090  
Db |||||  
QY 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTCTCCCTGCTGATCTCTGCGGATCTGCTG 793  
Db |||||  
QY 1091 GAACGCTCTGGGCTGTCGATCAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAA 1150  
Db |||||  
QY 794 GAACGCTGTGTGATCGTAGATAATATTTGACGCTTGGAACTTATAGACACTTTTGAAG 853  
Db |||||  
QY 1151 GAAATCAAGAAATCTCTGGAATACGTTTATAGGTACTTGGGACCCGAAAGAGCGTGGGA 1210  
Db |||||  
QY 854 GAAATCAAGGAAGCTCTTGATTTATGTTTACAGCAATTTGGAACG- - -AAGAGGAATTTGG 910  
Db |||||  
QY 1211 TGGGCAAGATGCAATCCTATTCTCTGATGTCGATGACACTGCCATGGGTCTTAGAATCCTG 1270  
Db |||||  
QY 911 TGGGCGAGACTAAATCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTGATTTGCTT 970  
Db |||||  
QY 1271 AGACTTCATGGATCAATGATCTTCHAGTGTCTTGGAATTTTCAAGACGAGAAAGGA 1330  
Db |||||  
QY 971 CGGCTGCATAGGTACAAATGATCTCCAGCCATTTTGGACAACTTCAAGATGCGCAATGGG 1030  
Db |||||  
QY 1331 GACTCTTTTGTCTTGGCGTCAAAACGCAATTTGGTGTGACCGATATCTTAACTTTAT 1390  
Db |||||  
QY 1031 AAATTCATTTGCTGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGGAATCTTTAT 1090  
Db |||||  
QY 1391 AGATGTTTCAAGATGATTTTTCCGGGAGAAAGATATGGAAGAGCTTAAGACCTTCACT 1450  
Db |||||  
QY 1091 AGAGCTTCCAGCTCGCATTTCCGGGAGANAACATTTCTGATGAAGCTTAAAGCTTCGCT 1150  
Db |||||  
QY 1451 ACAAATCATCTCCAAATGCTCTTGGCCAAAACAAACGCAATTTGATAGTGGGCTGTCAAG 1510  
Db |||||  
QY 1151 ACTAAATATTTGAGAGAAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAAACAA 1210  
Db |||||  
QY 1511 AAGGATCTTCTGAGAGGTGAGTAGTCTTAAAGTATCCGTGGCATAGAGTATGCCA 1570  
Db |||||  
QY 1211 CAAAACCTGAGCCAAAGAGATCAAAATACGCGCTGAAGACTTCTTGGCATGCGAGTTCCG 1270  
Db |||||  
QY 1571 AGATTGGAGCAAGAACTTTACATAGAGCAATTTGGATCAAAATGATGTTCTGGCTGGGGAAG 1630  
Db |||||  
QY 1271 AGAGTGAAGCAAGAGATACTGTCAAGTGTATCGCCCAAGATTTGCAACCATAGCAAAA 1330  
Db |||||  
QY 1631 ACTGTGTAAGATGCTATATGTGAGCAACGAAATAATTTGGAGCTGGCCAAATTTGGAC 1690

Db |||||  
QY 1331 TGGGTTTACAGTACCCTACGTGAACTGAAAGTTTTTTAGAGCTGGGAAATTAGAT 1390  
Db |||||  
QY 1691 TTCAATATGTTGCGAGCCTTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGGTGAGA 1750  
Db |||||  
QY 1391 TTCAACATTTATCCAGTCCATCCACCAAGAAAGAAATGAAGAAATGTACCAGCTGGTTAGA 1450  
Db |||||  
QY 1751 GAATCGGGAATCAATGATCTTACATTCACCGCCAGCGGCTGTGGAAATGTATTTCTCA 1810  
Db |||||  
QY 1451 GATTCGGGTTGCCACTATTTCACCTTCGCTGGGAGAGCGCTGGAATTTCTACTTTTA 1510  
Db |||||  
QY 1811 GTGGCGTTAGTATGTTTTCAGCCAGAAATTCGCTGTGTAGAAATTCGCTATGCAAGACT 1870  
Db |||||  
QY 1511 GTAGCGCGGAGCCTATGAAACCCAGTATGCCAATGCAAGTTCTCTTTACANAAGTG 1570  
Db |||||  
QY 1871 TCTTTCCTCGCAGTTATTCTTAGACGATTTTACGACCCACCGATCTCTGGAATGATCTT 1930  
Db |||||  
QY 1571 GCATGCTTCGACAGACTGTTCTGGACGATATGTATGACACTTATGGAACCCCTAGATGAATTG 1630  
Db |||||  
QY 1931 AAATGTTCTCTGAAAGCGGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGTTCCGGAT 1990  
Db |||||  
QY 1631 AAGCTATTCACTGAGGCTGTGAGAAATGGGACCTCTCTCTTTACAGAAAACCTTCCAGAC 1690  
Db |||||  
QY 1991 AATCAGTTCGAAAGTTTGTCTTTCAGGGCTGTACAAACACAGTGAATTCGAAATTCG 2050  
Db |||||  
QY 1691 TAT- - -ATGAAACTATGTTACCAAAATCTATTATGACATAGTTTACGAGGTGGCTTGGAG 1747  
Db |||||  
QY 2051 GGACTCAAGGAAACAGCGCGTGTATGCTGGGCTATCTTCGAAAAGTATGGGAGGGCTTG 2110  
Db |||||  
QY 1748 GCAGAGAAGGAAACAGGGCGTGAATTTGGTCAGCTTTTTCAGAAAGGATGGGAGGATTAT 1807  
Db |||||  
QY 2111 CTCGCATCTGTATACAAAGAGCCGAATGGTCCGAGCAAGATATGTGCCGACATTCNAAC 2170  
Db |||||  
QY 1808 CTTCTGGGTTTATATGAAGAAGCTGAATGGTTAGCTGTGCTGAGTATGTCCCTACCTTGGAC 1867  
Db |||||  
QY 2171 GAATATGTGAAAATGCCAAAGTGTCCATAGACACTTTCGACAGCTGCTACTAACTCAATC 2230  
Db |||||  
QY 1868 GAGTACATAAAGAAATGGAATTCATCTATCGGCCAAAGTATATCTTCTGTTGAGTGGAGTG 1927  
Db |||||  
QY 2231 TTTTTC- - -CTGGAGAAATTTACTTCTCGATTTACATTTTACAGCAAGTAGACTTTCGGTCC 2287  
Db |||||  
QY 1928 TTGATAATGGATGGGCAACTCTCTTTCGAAAGAGGCATTTAGAGAAAGTAGATTATCCAGGA 1987  
Db |||||  
QY 2288 AAATTTCTCG- - -ATCTTGTGCTTTTGAAGTGGAGCACTAATCAATGACACCAAGACT 2341  
Db |||||  
QY 1988 AGACGTGTTTCTCACAGAGCTGAATAGCCTCATTTCCCGCTGGCGGATGACACGAGACA 2047  
Db |||||  
QY 2342 TACCGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGTACATGAGGAGGAA 2401  
Db |||||  
QY 2048 TATAAAGCTGAGAAAGGCTCGTGGAATTTGGCGTCCAGCATTTGAATGTTTACATGAAGAAG 2107  
Db |||||  
QY 2402 AATCCCGAGTGCAAGAGGAAAGGCTCTAAAGTCAATGTTTATGTTATCATCGCAACGCA 2461  
Db |||||  
QY 2108 CATCTCGAATGTACAGAGGAGGCTCTCGATCACAATCTATAGCATTTCTGGAGCCGCG 2167  
Db |||||  
QY 2462 CTGAAGAAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCAATTTGTGTGAGA 2521  
Db |||||  
QY 2168 GTGAAGGAACTGCAAGAGAGAGTTTCTGAAGCCCG- - -ACGACGTCCTTCCCTCGCTGCAAG 2224  
Db |||||  
QY 2522 AGACTGCTGTCAACACTGCAAGAGTGTGACGCTGTTTTATATGTATGACAGAGATGGCTTT 2581  
Db |||||  
QY 2225 AGATGCTTTTCAGGAGACAGAGTGACGATGGTGATATTCAGGATGGAGATGGATTC 2284  
Db |||||  
QY 2582 GGTATCTCTGCAACAAAGAGATGAAAGACCACTGTCCAGCCGAACTCTTTTCGATCTCTGGCG 2641  
Db |||||  
QY 2285 GGTGTTTCCAAATTTAGAGTCAAGATCATATCAAAGAGTGTCTCATTTGAACCGCTGCCA 2344  
Db |||||  
QY 2642 TAGCA- - -TACTGATATTTATATATATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCA 2688  
Db |||||  
QY 2345 CTGTAATCAAAATGATTTGCAATTAATTTGAAATAATGTCAACTATGTCTTTCACAAAAA 2404  
Db |||||  
QY 2689 AAAAAAATAAAAAA 2705  
Db |||||

Db	2405	AAAAAAAAAAAAAAAA	2421	
RESULT 15				
US-09-900-797-45				
; Sequence 45, Application US/09900797				
; Patent No. 6645762				
; GENERAL INFORMATION:				
; APPLICANT: Chappell, Joseph				
; APPLICANT: No. 66457621, Joseph P.				
; APPLICANT: Starks, Courtney M.				
; APPLICANT: Manna, Kathleen R.				
; TITLE OF INVENTION: SYNTHASES				
; FILE REFERENCE: 07678-025001				
; CURRENT APPLICATION NUMBER: US/09/900,797				
; CURRENT FILING DATE: 2001-07-06				
; PRIOR APPLICATION NUMBER: US/09/398,395				
; PRIOR FILING DATE: 1999-09-17				
; PRIOR APPLICATION NUMBER: 60/130,628				
; PRIOR FILING DATE: 1999-04-22				
; PRIOR APPLICATION NUMBER: 60/150,262				
; PRIOR FILING DATE: 1999-08-23				
; NUMBER OF SEQ ID NOS: 58				
; SOFTWARE: FastSeq for Windows Version 3.0				
; SEQ ID NO 45				
; LENGTH: 2424				
; TYPE: DNA				
; ORGANISM: Abies grandis				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (2)...(2347)				
; OTHER INFORMATION: E-alpha-bisabolene synthase				
US-09-900-797-45				
Query Match				
Best Local Similarity 27.0%; Score 730.2; DB 4; Length 2424;				
Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;				
QY	434	AAATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATCCGTCATATGACGGCTCTGGT	493	
Db	134	ACTCCATCTGTTATGACACAGATGGGTAGCGGGGTGCCGCCATTGATGGCTCTGCT	193	
QY	494	GCACCCCAATTTCCCAACCGCTTCAATGGATTTCTGAACATCACTGCAGATGGCTCG	553	
Db	194	CGCCCGCAATTTCCCAACCGCTTCAATGGATTTCTGAACATCACTGCAGATGGCTCG	253	
QY	554	TGGGGTGAGGAGTGCATTTCTTGGCGTATGACAGAGTTTAAACACTCTCGCCCTGCCTC	613	
Db	254	TGGGGAAATTCAGTCCCACTTTCTGCTGTCGACCGCTCTTCTGCTCTCTTCTGTT	313	
QY	614	CTCACTCTCAAAATATGGAATAGGGCGACATTCAGATGCGAGAAAGGGTGTGATTTGG	673	
Db	314	CTTGTGCTCTTAAATGGAACGTTGGGGATCTGCAAGTAGAGCAGCGGAATTGAATTCAT	373	
QY	674	AGAAACACATGGAAGATGAGGACGAGCTGACATCACAG---GCCAAGTGGATTC	730	
Db	374	AAGAGCAATCTGGAATAGTAAAGGATGAACCCGATCAAGATAGCTTGGTAAACAGACTT	433	
QY	731	GAGGTGCTGTTCTCTCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGAATCTTCTTAT	790	
Db	434	GAGATCATATTTCTTCTCTGTTAAGAGACTCAATCTCTGCGCTCGGACTTCCCTAC	493	
QY	791	CACCTCCCTTTTCAATCTCCCAATTCACAAAGCGCCAGAAAGCTTCAAAAGATTC	850	
Db	494	GACCTGCTTATATACATCTGTTGCACTTAAACCGGAGGAAAGATTAGCAAACTTCA	553	
QY	851	CTCAATGTTCTTCAACCATCAGCGGCTGCTCTACTCTCTGGAGGGTGTGCAAGAT	910	
Db	554	AGGAGGAATTTATGCGGTTCCGTCGCCATTGTTGTTTCTTTAGAGGAATACAGAT	613	
QY	911	GTGGTGACTGGCAAGAGATCACAAAATCTTCAATCAAGACCGGATCATTTTAAAGTCC	970	
Db	614	ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTAAAGCTCA	673	
QY	971	CCTGCATCTACTGCTTGTGTTCTTATGCACTCAAAACAAACGATGCTCCTCCTCCTCTC	1030	
Db	674	CCTGCTTCTACTGCTGCTGTTTTCATGCAACAGAGACGGAATGCTTGAATCTTG	733	
QY	1031	AACCTCGTCTCAGCAAAATTTGGCGACTACGTTCTTGGCATTACCCACTTGTATTT	1090	
Db	734	AACAGTGTGATGATCAAGTTTGGAAATTTGTTCCCTGCTGTATCTCTGATCTGCTG	793	
QY	1091	GAAAGCTCTGGCTGCTGATACAGTTGAAAGCTTGGAAATCGATCGCTTATTTCAAGAA	1150	
Db	794	GAAAGCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	853	
QY	1151	GAAATCAAGAAATCTCTGATTTACGTTTATAGTTACTGGGACCGCAAGAGAGCGCTGGA	1210	
Db	854	GAAATCAAGAAAGCTCTTGAATTTATGTTTACAGGCAATGGAACG---AAAGAGGAAT	910	
QY	1211	TGGCAAGATGCAATCTTATCTGATGTCGATGACACTGCCATCGCTTATGAAATCTCTG	1270	
Db	911	TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACACTGCTTGGGATTTGATGCTT	970	
QY	1271	AGACTTTCATGGATACAAATGTTATCTTTCAGATGTTCTGGAGAAATTCAGACAGAGAA	1330	
Db	971	CGGCTGCATAGTACAAATGTTATCTCCAGCAATTTTGAACACTTCAAGATGCAATGG	1030	
QY	1331	GACTTCTTTTGGCTTGGCGTCAACCGCAATTTGGTGTGACCGATATCTTAACTTTAT	1390	
Db	1031	AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGATCTTTAT	1090	
QY	1391	AGATGTTTCAAGATGTTTTCGGGAGAAAAAGATAATGGAAGAGCTAAGACCTTCCT	1450	
Db	1091	AGAGCTTCCAGCTCGCATTTCCCGAGAAACATCTTGTATGAGTCAAAAGCTTCGCT	1150	
QY	1451	ACAAATCATCTCCAAATGCTCTTGGCAAAAACACGCAATTTGATTAAGTGGGTGTCAAG	1510	
Db	1151	ACTAAATATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAACAA	1210	
QY	1511	AAGATCTTCTGGAGAGTGGAGTATGCTATAAGTATCCGTGGCATAGAGTATGCCA	1570	
Db	1211	CAAAACCTGAGCAAGAGATCAAAATACGCGCTGAAGACTTCTTGGCATGCCAGTTCG	1270	
QY	1571	AGATGGAGCAAGAGTTTACATAGAGCAATTTGGATCAAAATGATGCTGTGGTGGGAAG	1630	
Db	1271	AGAGTGGAGCAAGAGATGATGTCAGGTGATCGCCAGATTTATGACGCAATAGCAAA	1330	
QY	1631	ACTGTGTAAGATGCTATATGACCAACGAAATATTTGGAGCTGGCCAAATTTGGAC	1690	
Db	1331	TGGTTTTCAAGCTACCTTACGTTGAACAAATGAAAGTTTGTAGAGCTGGGAAATTAG	1390	
QY	1691	TTCAATATGCTGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGGTGAGA	1750	
Db	1391	TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAGAAATGTTACCAAGCTGGTTA	1450	
QY	1751	GAAATCGGATTCAAATGATTTTACATTTACCCGCGAGGCTGTGGAAATGATTTCTCA	1810	
Db	1451	GATTCGGGTTGCGACTTATTCACCTTCTGCTCGGAGAGGCGCTGGAATTTCTACTTTA	1510	
QY	1811	GTGGCGTTAGTATGTTTGGAGCAGAAATTCGCTGCTTGTAGAAATGCTATGCAAGACT	1870	
Db	1511	GTAGCGCGGGGACCTATGAACCCAGTATGCCAAATGCAAGTTCTCTTTTACAAAGTG	1570	
QY	1871	TCTTGTCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCACCGATCTCTGGATGAT	1930	
Db	1571	GCATGCTTGCAGACTGTTCTGAGCAGATATGATGACACTTATGGAACCTTAGATGAATG	1630	
QY	1931	AAATGTTCTTGAAGGGTCCGAAGATGGGATATCTGTGCTGGATAGCGTTCGGGAT	1990	
Db	1631	AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCTTTTACAGAAACCTTCCAGAC	1690	
QY	1991	AATCAGTTGAAGCTTTGCTTCTTAGGGCTGTACAAACACAGTGAATGATTTGGAAGAT	2050	
Db	1691	TAT---ATGAACATATGTTACCAAAATCTATTATGACATAGTTTACAGAGGTGGCTGGAG	1747	

```
QY 2051 GGACTCAAGGACAAAGCGCTGATGTCTGGGCTATCTTCGAAAACTATGGGAGGCTTG 2110
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1748 GCAGAGAGAGACAGGGGGTGAATTGGTCAGCTTTTTCAGAAAGGATGGGAGGATTAT 1807
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2111 CTCGCATCGTATATACCAAGAAGCCGAATGGTCGGCAGCAAAAGTATGTGCCGACATTCAAC 2170
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1808 CTTCTGGGTATTATGCAAGAAGCTGAATGGTTAGCTGCTGAGTATGTGCTACCTTGGAC 1867
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2171 GAATATGTGGAANAATGCCAAAGTGTCATAGACACTTGGACAGTCGTACTAAACTCAATC 2230
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1868 GAGTACATAAAGAATGGAATCAATCTATCGGCCAACGTATATCTTCTGTTGAGTGGAGTG 1927
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2231 TTTTTCAC---CTGGAGAAATTACTTCTCGATTATACATTTTACGCAAGTAGACCTTCGGTCC 2287
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1928 TTGATAAATGGATGGGCACTCTCTTCGCAAGGGCAATTAGAGAAAGTAGATTATCCAGGA 1987
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2288 AAATTTCTGC-----ATCTTGTCCTTTGACTGGACGACTAATCAATGACACCAAGACT 2341
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1988 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCCGCTGGCGGATGACACGAAGACA 2047
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2342 TACGAGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGTACATGAGGGAA 2401
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2048 TATAAAGCTGAGAAGCTCGTGGAGAAATGGCGTCCAGCAATTGAATGTTACATGAAAGAC 2107
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2402 AATCCGAGTGCACAGAGCAAGAGCTCTAAGTCATGTTTATGGTATCATCGACACGCA 2461
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2108 CATCCTGAATGTACAGAGGAAGAGGCTCTCGATCAATCTATAGCAATCTGGAGCCGGCG 2167
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2462 CTGAAGGAATTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCAATTGTGTGTGAGA 2521
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2168 GTGAGGAAGTGCACAAGAGATTCTGAAGCCG---ACGAGTCCCATTCGCCTGCAAG 2224
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2522 AGACTGCTGTTCAACACCTGCAAGAGTGAATGCAGCTGTTTTATATGTAAGAGATGGCTTT 2581
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2225 AAGATGCTTTTCGAGGAGACAAGAGTGCAGTGGTGATATTCAGGATGGAGATGGATTTC 2284
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2582 GGTATCTCTGACAAAGAGATGAAGACCATGTCAGCCGAACTCTTTTCGATCCTGTGGCG 2641
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2285 GGTGTTTCCAAATTAGAAGTCAAAGATCATATCAAAGAGTGTCTCATTTGAACCGCTGCCA 2344
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2642 TAGCA-----TACTGATATTATATATATATTATTTCAATTCAAAAAAA 2688
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2345 CTGTAATCAAATAGTTGCAATAATAATTGAATAATGTCAACTATGTTTCACAAAAAA 2404
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2689 AAAAAAAAAAAAAAA 2705
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2405 AAAAAAAAAAAAAAA 2421
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: August 24, 2004, 21:21:01  
Job time : 187.026 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 15:24:18 ; Search time 1182.41 Seconds  
(without alignments)  
11237.470 Million cell updates/sec

Title: US-10-041-007-1  
Perfect score: 2705  
Sequence: 1 atattgcctaaccctgcac.....aaaaaaaaaaaaaaaaaaaaa 2705

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3228839 seqs, 245606551 residues

Total number of hits satisfying chosen parameters: 6457678

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgm2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:
- 2: /cgm2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:
- 3: /cgm2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:
- 4: /cgm2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:
- 5: /cgm2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:
- 6: /cgm2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:
- 7: /cgm2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:
- 8: /cgm2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:
- 9: /cgm2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:
- 10: /cgm2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:
- 11: /cgm2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:
- 12: /cgm2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:
- 13: /cgm2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:
- 14: /cgm2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:
- 15: /cgm2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:
- 16: /cgm2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:
- 17: /cgm2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:
- 18: /cgm2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:
- 19: /cgm2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2705	100.0	2705	12	US-10-041-018-397
2	2705	100.0	2705	14	US-10-041-007-1
3	2622	96.9	2622	14	US-10-041-007-32
4	2442	90.3	2445	14	US-10-041-007-34
5	2385	88.2	2388	14	US-10-041-007-36
6	2241	82.8	2241	14	US-10-041-007-38
7	1091.4	40.3	2861	9	US-09-887-586A-55
8	1091.4	40.3	2861	9	US-09-903-012-55
9	1091.4	40.3	2861	13	US-09-900-797-55
10	1091.4	40.3	2861	12	US-10-041-018-365
11	1091.4	40.3	2861	13	US-09-893-820-55
12	1091.4	40.3	2861	14	US-10-041-007-3
13	730.2	27.0	2424	9	US-09-887-586A-45
14	730.2	27.0	2424	9	US-09-903-012-45

15	730.2	27.0	2424	10	US-09-900-797-45	Sequence 45, Appl
16	730.2	27.0	2424	13	US-09-893-820-45	Sequence 45, Appl
17	730.2	27.0	2424	14	US-10-041-007-14	Sequence 14, Appl
18	730.2	27.0	2424	15	US-10-025-145A-15	Sequence 15, Appl
19	730.2	27.0	2528	14	US-10-041-007-13	Sequence 13, Appl
20	642.2	23.7	2700	9	US-09-887-586A-43	Sequence 43, Appl
21	642.2	23.7	2700	9	US-09-903-012-43	Sequence 43, Appl
22	642.2	23.7	2700	10	US-09-900-797-43	Sequence 43, Appl
23	642.2	23.7	2700	12	US-10-041-018-363	Sequence 363, App
24	642.2	23.7	2700	13	US-09-893-820-43	Sequence 43, Appl
25	450.4	16.7	2186	15	US-10-025-145A-66	Sequence 66, Appl
26	416	15.4	1890	15	US-10-025-145A-77	Sequence 77, Appl
27	412.2	15.2	2205	15	US-10-025-145A-31	Sequence 31, Appl
28	411.2	15.2	1865	9	US-09-887-586A-47	Sequence 47, Appl
29	411.2	15.2	1865	9	US-09-903-012-47	Sequence 47, Appl
30	411.2	15.2	1865	10	US-09-900-797-47	Sequence 47, Appl
31	411.2	15.2	1865	13	US-09-893-820-47	Sequence 47, Appl
32	411.2	15.2	1865	14	US-10-041-007-17	Sequence 17, Appl
33	410.6	15.2	2196	9	US-09-887-586A-29	Sequence 29, Appl
34	410.6	15.2	2196	9	US-09-903-012-29	Sequence 29, Appl
35	410.6	15.2	2196	10	US-09-900-797-29	Sequence 29, Appl
36	410.6	15.2	2196	13	US-09-893-820-29	Sequence 29, Appl
37	410.6	15.2	2196	14	US-10-041-007-25	Sequence 25, Appl
38	410.6	15.2	2196	15	US-10-025-145A-1	Sequence 1, Appl
39	409.4	15.1	2460	9	US-09-371-307-2	Sequence 2, Appl
40	409.4	15.1	2460	16	US-10-401-321-2	Sequence 2, Appl
41	407.2	15.1	2089	9	US-09-887-586A-57	Sequence 57, Appl
42	407.2	15.1	2089	9	US-09-903-012-57	Sequence 57, Appl
43	407.2	15.1	2089	10	US-09-900-797-57	Sequence 57, Appl
44	407.2	15.1	2089	13	US-09-893-820-57	Sequence 57, Appl
45	407.2	15.1	2089	14	US-10-041-007-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-10-041-018-397  
; Sequence 397, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041.018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 397  
; LENGTH: 2705  
; TYPE: DNA  
; ORGANISM: Ginkgo biloba  
US-10-041-018-397

Query Match	100.0%	Score 2705;	DB 12;	Length 2705;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2705;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATATTTGCTTAAACCTGCACAAATGGCTGGGGTCTTTTGGAAATCTGCTTGTCTACT	60	
Db	1	ATATTTGCTTAAACCTGCACAAATGGCTGGGGTCTTTTGGAAATCTGCTTGTCTACT	60	
Qy	61	GCAACTCTCTCCAAAGTTCCTTCGGCAATCCACATAATCTTCTTCTTTTCAAA	120	
Db	61	GCAACTCTCTCCAAAGTTCCTTCGGCAATCCACATAATCTTCTTCTTTTCAAA	120	
Qy	121	GAGATCTCTATTGGATTTAATGACAGCACTGGTCCGTTCTCATTAGAGCTGAGATG	180	
Db	121	GAGATCTCTATTGGATTTAATGACAGCACTGGTCCGTTCTCATTAGAGCTGAGATG	180	



QY 181 GAAATGTCGGGATTCATGCTCAGCTGCAGACTCGTCCAGATCAGCTTCCACAGGA 240  
 Db 181 GAAATGTCGGGATTCATGCTCAGCTGCAGACTCGTCCAGATCAGCTTCCACAGGA 240  
 QY 241 GGAACGCTTTGTGTCGAGACTTAATTCGGGATTTATCATCCAGCTGTCTGGAAAGCGATTT 300  
 Db 241 GGAACGCTTTGTGTCGAGACTTAATTCGGGATTTATCATCCAGCTGTCTGGAAAGCGATTT 300  
 QY 301 CATCGACTCTCTAAACATCCCTTAATTCGACGCGACATCGAAATCAAGCGTCGATGAGAC 360  
 Db 301 CATCGACTCTCTAAACATCCCTTAATTCGACGCGACATCGAAATCAAGCGTCGATGAGAC 360  
 QY 361 AATCAATAAAGAAATCCAGACATTCGTAAGGAAATCCAGTGCATGTTTCAGTCCATGGG 420  
 Db 361 AATCAATAAAGAAATCCAGACATTCGTAAGGAAATCCAGTGCATGTTTCAGTCCATGGG 420  
 QY 421 CGACGGTGAAGAAATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATCCGTCAAT 480  
 Db 421 CGACGGTGAAGAAATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATCCGTCAAT 480  
 QY 481 TGAACGCTCTGTGCACCCCAATTTCCCAACGCTTCAATGGATCTGAAACAACTCAACT 540  
 Db 481 TGAACGCTCTGTGCACCCCAATTTCCCAACGCTTCAATGGATCTGAAACAACTCAACT 540  
 QY 541 GCAGATGCTCTGTGGGTGAGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACAC 600  
 Db 541 GCAGATGCTCTGTGGGTGAGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACAC 600  
 QY 601 TCTCGCTCGCTCCTCACTCTCAAAATATGGAATTAAGGCGGACATTTCAAGTCGAGAAAG 660  
 Db 601 TCTCGCTCGCTCCTCACTCTCAAAATATGGAATTAAGGCGGACATTTCAAGTCGAGAAAG 660  
 QY 661 GGTGAGTTTGTGAGAAACACATGGAAGAAATGAAGACGAGCTGACAAATCAGAGGCC 720  
 Db 661 GGTGAGTTTGTGAGAAACACATGGAAGAAATGAAGACGAGCTGACAAATCAGAGGCC 720  
 QY 721 AAGTGATTCGAGGTGCTGTTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGA 780  
 Db 721 AAGTGATTCGAGGTGCTGTTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGA 780  
 QY 781 TCTTCCTTATCAGCTCCCTTCACTCCCAATCCACCAAGGCGCCAGAAAGCTTCA 840  
 Db 781 TCTTCCTTATCAGCTCCCTTCACTCCCAATCCACCAAGGCGCCAGAAAGCTTCA 840  
 QY 841 AAGATTCCTCCTCAATGTTCTTCAATCAACCATCAGAGCGGCTTCTCTCTCGAGGG 900  
 Db 841 AAGATTCCTCCTCAATGTTCTTCAATCAACCATCAGAGCGGCTTCTCTCTCGAGGG 900  
 QY 901 TTTGCAAGATGTTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGCGGATCAAT 960  
 Db 901 TTTGCAAGATGTTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGCGGATCAAT 960  
 QY 961 TTTAAGCTCCCTGCACTACTGCTGTGCTTCAATGACACACTCAAAACAAAGCGCT 1020  
 Db 961 TTTAAGCTCCCTGCACTACTGCTGTGCTTCAATGACACACTCAAAACAAAGCGCT 1020  
 QY 1021 CCACTTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTACGTTCTTGGCCATTAACCACT 1080  
 Db 1021 CCACTTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTACGTTCTTGGCCATTAACCACT 1080  
 QY 1081 TGATCTAATTTGAAGCCTCTGGGCTGTCGATACAGTTGAAGCCTTGGGAATCGATCGCTA 1140  
 Db 1081 TGATCTAATTTGAAGCCTCTGGGCTGTCGATACAGTTGAAGCCTTGGGAATCGATCGCTA 1140  
 QY 1141 TTTCAAGAAAGAAATCAAGAAATCTCTGGATTAAGTTATAGGTACTTGGGACGCGGAAAG 1200  
 Db 1141 TTTCAAGAAAGAAATCAAGAAATCTCTGGATTAAGTTATAGGTACTTGGGACGCGGAAAG 1200  
 QY 1201 AGGCGTGGGATGGCAAGATGCAATCTTAATTCCTGATGTCGATGACACTGCGCATGGGTCT 1260  
 Db 1201 AGGCGTGGGATGGCAAGATGCAATCTTAATTCCTGATGTCGATGACACTGCGCATGGGTCT 1260  
 QY 1261 TAGAATCCTGAGACTTTCATGGATACATGATATCTTCAGATGTTCTTGGGAATTTTCAGAGA 1320

Db 1361 TAGAATCCTGAGACTTTCATGGATCAATGTATCTTTCAGATGTTCTTGGGAATTTTCAGAGA 1320  
 QY 1321 CGAAGAGGAGACTTCTTTTCTGCTTTCGCGGTCAACGCAAAATTTGGTGTGACCGATAATCT 1380  
 Db 1321 CGAAGAGGAGACTTCTTTTCTGCTTTCGCGGTCAACGCAAAATTTGGTGTGACCGATAATCT 1380  
 QY 1381 TAACTTTTATAGATGTTTCACAAGATGATGTTTTCGCGGAGAAAGATTAATGGAAAGCTAA 1440  
 Db 1381 TAACTTTTATAGATGTTTCACAAGATGATGTTTTCGCGGAGAAAGATTAATGGAAAGCTAA 1440  
 QY 1441 GACCTTCACTCAAAATCATCTCCAAATGCTTTCGCAAAACACGCAATTTGATAAGTG 1500  
 Db 1441 GACCTTCACTCAAAATCATCTCCAAATGCTTTCGCAAAACACGCAATTTGATAAGTG 1500  
 QY 1501 GGCTGTCAAGAAGGATCTTCTGAGAGGTGAGTATGCTATAAAGTATCCGTGGCATAG 1560  
 Db 1501 GGCTGTCAAGAAGGATCTTCTGAGAGGTGAGTATGCTATAAAGTATCCGTGGCATAG 1560  
 QY 1561 AAGTATGCAAGATTTGGAGGCAAGAAATTAATAGAGCAATTTGGATCAAAATGATGCTG 1620  
 Db 1561 AAGTATGCAAGATTTGGAGGCAAGAAATTAATAGAGCAATTTGGATCAAAATGATGCTG 1620  
 QY 1621 GCTGGGAGAGCTGTGATATAGATCTATATGAGCAACGAAATATTTGGAGCTGGC 1680  
 Db 1621 GCTGGGAGAGCTGTGATATAGATCTATATGAGCAACGAAATATTTGGAGCTGGC 1680  
 QY 1681 CAAATTTGACCTTCAATATGTTGTCAGGCCCTTACACAAAGGAGACTCAACATTTGTCAG 1740  
 Db 1681 CAAATTTGACCTTCAATATGTTGTCAGGCCCTTACACAAAGGAGACTCAACATTTGTCAG 1740  
 QY 1741 CTGGTGGAGAAATCGGGAATTCATGATTTACATTTACCCCGCAGCGCTGTGGAAT 1800  
 Db 1741 CTGGTGGAGAAATCGGGAATTCATGATTTACATTTACCCCGCAGCGCTGTGGAAT 1800  
 QY 1801 GTATTTCTCAGTGGCGTTAGTATGTTGAGCGCAGAAATTCGCTGTGTAGAAATTCCTTA 1860  
 Db 1801 GTATTTCTCAGTGGCGTTAGTATGTTGAGCGCAGAAATTCGCTGTGTAGAAATTCCTTA 1860  
 QY 1861 TGCCAAGACTTCTTGCCTCGCAGTTATTTAGAGCATTTTACGACACCCACCGATCTCT 1920  
 Db 1861 TGCCAAGACTTCTTGCCTCGCAGTTATTTAGAGCATTTTACGACACCCACCGATCTCT 1920  
 QY 1921 GATGATCTTAATTTGTTCTGAGCGCTCGAGATGATCTCTGTGTGATAG 1980  
 Db 1921 GATGATCTTAATTTGTTCTGAGCGCTCGAGATGATCTCTGTGTGATAG 1980  
 QY 1981 CGTTCCGGGATAATCAGTTGAAAGTTTGTCTTCTAGGGCTGTACAAACACAGTGAATGAT 2040  
 Db 1981 CGTTCCGGGATAATCAGTTGAAAGTTTGTCTTCTAGGGCTGTACAAACACAGTGAATGAT 2040  
 QY 2041 TGGAAAGATGAGCTCAAGGCAAGCGGTGATGCTGGGCTATCTTCGAAAGATG 2100  
 Db 2041 TGGAAAGATGAGCTCAAGGCAAGCGGTGATGCTGGGCTATCTTCGAAAGATG 2100  
 QY 2101 GAGGCTTGTCTCGCATCGTATACCAAGAGCGGAATGCTCGGAGCAAAAGTATGTC 2160  
 Db 2101 GAGGCTTGTCTCGCATCGTATACCAAGAGCGGAATGCTCGGAGCAAAAGTATGTC 2160  
 QY 2161 GACATTTCAACGAATATGCGAAATGCGCAAGTGTCCATAGCACTTGGCAGCTGCTACT 2220  
 Db 2161 GACATTTCAACGAATATGCGAAATGCGCAAGTGTCCATAGCACTTGGCAGCTGCTACT 2220  
 QY 2221 AAATCAATCTTTTCACTGGGAAATTTACTTCTGATTAATTTTACGCAAGTAGACCT 2280  
 Db 2221 AAATCAATCTTTTCACTGGGAAATTTACTTCTGATTAATTTTACGCAAGTAGACCT 2280  
 QY 2281 TCCGTCCAAATTTCTGATCTTGTGCTTTGACTTGGACGACTAATCAATGACCAAGAC 2340  
 Db 2281 TCCGTCCAAATTTCTGATCTTGTGCTTTGACTTGGACGACTAATCAATGACCAAGAC 2340  
 QY 2341 TTTACCGCGCGAGAAACCGTGTGATGTTGTTCCAGCGTACAGTCTACATGAGGA 2400



Db 1391 TAACTTTATAGATGTTTCAAGATATGTTTTCGGGAGAAAGATAATGGAAGAGCTAA 1440  
Qy 1441 GACCTTCACTCAAAATCATCTCCAAATGCTCTTGGCCAAAACCAACGCAATTTGTAAGTG 1500  
Db 1441 GACCTTCACTCAAAATCATCTCCAAATGCTCTTGGCCAAAACCAACGCAATTTGTAAGTG 1500  
Qy 1501 GCGTGTCAAGAGGATCTTCTGGAGAGTGAGTATCTATTAAGTATCCGTGGCATAG 1560  
Db 1501 GCGTGTCAAGAGGATCTTCTGGAGAGTGAGTATCTATTAAGTATCCGTGGCATAG 1560  
Qy 1561 AAGTATGCCAAGATTTGGAGGCAAGAGTTACATAGACAAATTTGGATCAATATGATGTCG 1620  
Db 1561 AAGTATGCCAAGATTTGGAGGCAAGAGTTACATAGACAAATTTGGATCAATATGATGTCG 1620  
Qy 1621 GCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAATAATTTGGAGCTGCG 1680  
Db 1621 GCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAATAATTTGGAGCTGCG 1680  
Qy 1681 CAAATTTGGACTTCAATATAGTGTGAGGCTTACACAAAGAGAGACTCAACACATTTGTGAG 1740  
Db 1681 CAAATTTGGACTTCAATATAGTGTGAGGCTTACACAAAGAGAGACTCAACACATTTGTGAG 1740  
Qy 1741 CTGGTGGAGAAATCGGATTCATATGATCTTACATTCACCCGCCAGCGCTGTGGAAT 1800  
Db 1741 CTGGTGGAGAAATCGGATTCATATGATCTTACATTCACCCGCCAGCGCTGTGGAAT 1800  
Qy 1801 GTATTTCTCAGTGGCGGTTAGTATGTTTTCAGCCAGAAATTCGCTGCTGTAGAAATTCGCTTA 1860  
Db 1801 GTATTTCTCAGTGGCGGTTAGTATGTTTTCAGCCAGAAATTCGCTGCTGTAGAAATTCGCTTA 1860  
Qy 1861 TGCCAAAGACTTCTTGCTCGAGTATTTCTAGACGATCTTTACGACCAACCGATCTCT 1920  
Db 1861 TGCCAAAGACTTCTTGCTCGAGTATTTCTAGACGATCTTTACGACCAACCGATCTCT 1920  
Qy 1921 GGATGATCTTAAATTTGCTTGAAGCGGTGCGAAGATGGATATCTCTGCTGCTGATAG 1980  
Db 1921 GGATGATCTTAAATTTGCTTGAAGCGGTGCGAAGATGGATATCTCTGCTGCTGATAG 1980  
Qy 1981 CGTTCCGGGATAATCAGTTGAAAGTTGCTTCTTAGGGCTGTACAAACAGTGAATGGATT 2040  
Db 1981 CGTTCCGGGATAATCAGTTGAAAGTTGCTTCTTAGGGCTGTACAAACAGTGAATGGATT 2040  
Qy 2041 TGGAAAGAGTGGACTCAAGGAAACAGCCGATGATGCTGGGCTATCTTCGAAAGATAG 2100  
Db 2041 TGGAAAGAGTGGACTCAAGGAAACAGCCGATGATGCTGGGCTATCTTCGAAAGATAG 2100  
Qy 2101 GGAGGCTTCTCGCATCGTATACCAAGAGCCGATGCTGGCAGCAAGATGATGTGCC 2160  
Db 2101 GGAGGCTTCTCGCATCGTATACCAAGAGCCGATGCTGGCAGCAAGATGATGTGCC 2160  
Qy 2161 GACATTCACGAAATATGTGAAATATGCAAGTGTCCATAGACTTTGGCGACGTGCTACT 2220  
Db 2161 GACATTCACGAAATATGTGAAATATGCAAGTGTCCATAGACTTTGGCGACGTGCTACT 2220  
Qy 2221 AAATCAATCTTTTCTGAGGATTTACTTCTGATTTATCATTTTACAGCAAGTAGACCT 2280  
Db 2221 AAATCAATCTTTTCTGAGGATTTACTTCTGATTTATCATTTTACAGCAAGTAGACCT 2280  
Qy 2281 TCGGTCCAAATTTCTGATCTTGTGCTTTGACTGGACGACTAATCAATGACCAAGAC 2340  
Db 2281 TCGGTCCAAATTTCTGATCTTGTGCTTTGACTGGACGACTAATCAATGACCAAGAC 2340  
Qy 2341 TTACAGGCCGAGAGAAACCGTGTGTGAATTTGGTTTCAGCGGTACAGTCTACATGAGGA 2400  
Db 2341 TTACAGGCCGAGAGAAACCGTGTGTGAATTTGGTTTCAGCGGTACAGTCTACATGAGGA 2400  
Qy 2401 AAATCCGAGTGACAGAGGAAAGAGCTCTAAGTCAATGTTTATGGTATCATCGCAACAGC 2460  
Db 2401 AAATCCGAGTGACAGAGGAAAGAGCTCTAAGTCAATGTTTATGGTATCATCGCAACAGC 2460  
Qy 2461 ACTGAAGGAAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAG 2520  
Db 2461 ACTGAAGGAAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAG 2520

RESULT 3

US-10-041-007-32  
; Sequence 32, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081051  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 2622  
; TYPE: DNA  
; ORGANISM: Ginkgo biloba  
US-10-041-007-32

Query Match 96.9%; Score 2622; DB 14; Length 2622;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 ATGGCTGGGGTGTCTTTTGCATAATCTGCTCTCACTGCAACTCTCTCTCAAAAGTTCCC 82  
Db 1 ATGGCTGGGGTGTCTTTTGCATAATCTGCTCTCACTGCAACTCTCTCTCAAAAGTTCCC 60  
Qy 83 TTCCGGCAATCCACTAATATCTTATCTCTTTCACAGAGATCCCTCATTTGGATTAAAT 142  
Db 61 TTCCGGCAATCCACTAATATCTTATCTCTTTTCAAGAGATCCCTCATTTGGATTAAAT 120  
Qy 143 GCACAGCACTGGCTCGGTTCTCACTTAAGGCTGAGATGGAATTTGTGCGGATTCATGCC 202  
Db 121 GCACAGCACTGGCTCGGTTCTCACTTAAGGCTGAGATGGAATTTGTGCGGATTCATGCC 180  
Qy 203 TCAGCTGTCAGAGACTCGTCCAGATCAGCTTCCAGAGGAAACGCTTTGTGTCGAGACTT 262  
Db 181 TCAGCTGTCAGAGACTCGTCCAGATCAGCTTCCAGAGGAAACGCTTTGTGTCGAGACTT 240  
Qy 263 AATGCGGATTTATCATCCAGCTGTCTGMAAGGAGATTTTCATCGACTCTCTAAACATCCCT 322  
Db 241 AATGCGGATTTATCATCCAGCTGTCTGMAAGGAGATTTTCATCGACTCTCTAAACATCCCT 300  
Qy 323 AATTTCCACCGGACATCGAAATCAAGCGTCAAGAGATTCATTAAGAAATCCAGACA 382  
Db 301 AATTTCCACCGGACATCGAAATCAAGCGTCAAGAGATTCATTAAGAAATCCAGACA 360  
Qy 383 TTGGTGAAGAAATCCAGTGCATGTTTCACTGCTGAGGAGGATTTTCATCGACTCTCTAAACATCCCT 442  
Db 361 TTGGTGAAGAAATCCAGTGCATGTTTCACTGCTGAGGAGGATTTTCATCGACTCTCTAAACATCCCT 420  
Qy 443 GCATATCATACAGCTTGGGTGGCAAGAAATTCGCTCAATTTGACGCTCTGTFGCACCCCAA 502

Db 421 GCATATGATACAGCTTGGGTGGCAAGAAATTCGGTCAATTGACGGCTCTGGTGCACCCCAA 480  
Qy 503 TTTCCCAACAGCTTCAATGATTCGAACAATCAACTGCGACAGATGGCTCGTGGGTGAG 562  
Db 481 TTTCCCAACAGCTTCAATGATTCGAACAATCAACTGCGACAGATGGCTCGTGGGTGAG 540  
Qy 563 GAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCGCTCCCTCACTCTC 622  
Db 541 GAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCGCTCCCTCACTCTC 600  
Qy 623 AAAATATGGAATAAGGCGACATTCAGTGCAGAAAAGGGTTCAGTTTGTGAGAAAACAC 682  
Db 601 AAAATATGGAATAAGGCGACATTCAGTGCAGAAAAGGGTTCAGTTTGTGAGAAAACAC 660  
Qy 683 ATGGAAGAAATGAAGGACGAGCTGACAATCAACGCGCAAGTGGATTCGAGGTCGTGTTT 742  
Db 661 ATGGAAGAAATGAAGGACGAGCTGACAATCAACGCGCAAGTGGATTCGAGGTCGTGTTT 720  
Qy 743 CCTGCAATGTTAGATGAAGCAAAAAGCTTGGGATTCGATCTCTTATCACTCCCTTTC 802  
Db 721 CCTGCAATGTTAGATGAAGCAAAAAGCTTGGGATTCGATCTCTTATCACTCCCTTTC 780  
Qy 803 ATCTCCCAATCCACAAAAGCGCAGAAAAGCTTCAAAAAGATTCCTCCCTCAATGTTCTT 862  
Db 781 ATCTCCCAATCCACAAAAGCGCAGAAAAGCTTCAAAAAGATTCCTCCCTCAATGTTCTT 840  
Qy 863 CATAACCATCAGACGGGTGCTCTACTCTGAGAGGTTTGCAAGATGTTGGACTGG 922  
Db 841 CATAACCATCAGACGGGTGCTCTACTCTGAGAGGTTTGCAAGATGTTGGACTGG 900  
Qy 923 CAAGAGATCACAATCTTCAATCAAGAGACGATCATTTTAAAGCTCCCTGCACTACT 982  
Db 901 CAAGAGATCACAATCTTCAATCAAGAGACGATCATTTTAAAGCTCCCTGCACTACT 960  
Qy 983 GCTTGTGCTTTCATGCACTCAAAAACAAACGATGCTCCCACTTTCTCAACTTCGTGCTC 1042  
Db 961 GCTTGTGCTTTCATGCACTCAAAAACAAACGATGCTCCCACTTTCTCAACTTCGTGCTC 1020  
Qy 1043 AGCAAAATTTGGGACTACGTTCTTGGCTTCCATTAACCACTGATCTATTTGAAAGCTCTGG 1102  
Db 1021 AGCAAAATTTGGGACTACGTTCTTGGCTTCCATTAACCACTGATCTATTTGAAAGCTCTGG 1080  
Qy 1103 GCTGTCGATACGTTGAAAGCTTGGGATCGATCGCTATTTCAAGAAAGAAATCAAGAA 1162  
Db 1081 GCTGTCGATACGTTGAAAGCTTGGGATCGATCGCTATTTCAAGAAAGAAATCAAGAA 1140  
Qy 1163 TCTCTGATTAACGTTTATAGGTACTGGGACCGCGAAAGCGCTGGGATGGGCAAGATGC 1222  
Db 1141 TCTCTGATTAACGTTTATAGGTACTGGGACCGCGAAAGCGCTGGGATGGGCAAGATGC 1200  
Qy 1223 AATCCTATTCCTGATGCGATGACATGCGCATGGTCTTAGAATTCCTGAGACTTCAAGGA 1282  
Db 1201 AATCCTATTCCTGATGCGATGACATGCGCATGGTCTTAGAATTCCTGAGACTTCAAGGA 1260  
Qy 1283 TACAATGATATTCAGATGTTCTGGAGAAATTCAGAGACGAGAAAGGAGACTTCTTTTGC 1342  
Db 1261 TACAATGATATTCAGATGTTCTGGAGAAATTCAGAGACGAGAAAGGAGACTTCTTTTGC 1320  
Qy 1343 TTTGGCGGTCAAAACGCAAAATTTGGTGTGACCGATAATCTTAAACCTTTATAGATGTTCAAA 1402  
Db 1321 TTTGGCGGTCAAAACGCAAAATTTGGTGTGACCGATAATCTTAAACCTTTATAGATGTTCAAA 1380  
Qy 1403 GTATGTTTTCGGGAGAAAAGATAATGGAAGAAGCTTAAGACCTTCACTACAAATCACTTC 1462  
Db 1381 GTATGTTTTCGGGAGAAAAGATAATGGAAGAAGCTTAAGACCTTCACTACAAATCACTTC 1440  
Qy 1463 CAAAATGCTCTTGGCAAAAACAGCAATTTGATAAGTGGGCTGTCAAGAGGATCTTCCT 1522  
Db 1441 CAAAATGCTCTTGGCAAAAACAGCAATTTGATAAGTGGGCTGTCAAGAGGATCTTCCT 1500  
Qy 1523 GGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAGTATGCCAAGATTTGGAGCA 1582  
Db 1501 GGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAGTATGCCAAGATTTGGAGCA 1560

Qy 1583 AGAAGTTACATAGAGCAATTTTGGATCAAAATGATGTCTGGCTGGGAAGACATGTGTATAAG 1642  
Db 1561 AGAAGTTACATAGAGCAATTTTGGATCAAAATGATGTCTGGCTGGGAAGACATGTGTATAAG 1620  
Qy 1643 ATCTTATATGTGAGCAACGAAAATATTTGGAGCTGGCCAAAATTTGGAGCTTCAATATGGTG 1702  
Db 1621 ATCTTATATGTGAGCAACGAAAATATTTGGAGCTGGCCAAAATTTGGAGCTTCAATATGGTG 1680  
Qy 1703 CAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCTAGCTGGTGGAGAGAAATCGGATTC 1762  
Db 1681 CAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCTAGCTGGTGGAGAGAAATCGGATTC 1740  
Qy 1763 AATGATCTTACATTTCACTCCCGCCAGCGCTGTGCGAAATGTATTTCTCAGTGGCGGTTAGT 1822  
Db 1741 AATGATCTTACATTTCACTCCCGCCAGCGCTGTGCGAAATGTATTTCTCAGTGGCGGTTAGT 1800  
Qy 1823 ATGTTTGGAGCCAGAAATTCGCTGCTGTGTAGAAATTCCTATGCCAAGACTTCTTGGCTCGCA 1882  
Db 1801 ATGTTTGGAGCCAGAAATTCGCTGCTGTGTAGAAATTCCTATGCCAAGACTTCTTGGCTCGCA 1860  
Qy 1883 GTTATTTCTAGACGATCTTTACGACACCCACGATCTCTGGATGATCTTAAATTTCTCT 1942  
Db 1861 GTTATTTCTAGACGATCTTTACGACACCCACGATCTCTGGATGATCTTAAATTTCTCT 1920  
Qy 1943 GAAGCGTCCGAAGATGGGATATCTCTGTCTGATAGCGTTCCGGATTAATCAGTTGAAA 2002  
Db 1921 GAAGCGTCCGAAGATGGGATATCTCTGTCTGATAGCGTTCCGGATTAATCAGTTGAAA 1980  
Qy 2003 GTTGTCTTCTAGCGGTGTACAAACAGTGAATTTGGAAAAGATGCACTCAAGAA 2062  
Db 1981 GTTGTCTTCTAGCGGTGTACAAACAGTGAATTTGGAAAAGATGCACTCAAGAA 2040  
Qy 2063 CAAAGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGAGGGCTTGTCTCGCATCGTAT 2122  
Db 2041 CAAAGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGAGGGCTTGTCTCGCATCGTAT 2100  
Qy 2123 ACCAAAGAGCCGAATGGTCCGAGCAAAAGTATGTGCCGACATTCACGAATATGTGGAA 2182  
Db 2101 ACCAAAGAGCCGAATGGTCCGAGCAAAAGTATGTGCCGACATTCACGAATATGTGGAA 2160  
Qy 2183 AATGCCAAAGTGTCCATAGCACCTTGCAGACAGTCTGTAAACTCAATCTTTTTCACCTGGA 2242  
Db 2161 AATGCCAAAGTGTCCATAGCACCTTGCAGACAGTCTGTAAACTCAATCTTTTTCACCTGGA 2220  
Qy 2243 GAATTAATCTCTGATTAACATTTTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCACTTT 2302  
Db 2221 GAATTAATCTCTGATTAACATTTTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCACTTT 2280  
Qy 2303 GTGTCTTTGACTGGAGCACTAATCAATGACACCAAGACTTTACAGGCCGAGAGAAACCGT 2362  
Db 2281 GTGTCTTTGACTGGAGCACTAATCAATGACACCAAGACTTTACAGGCCGAGAGAAACCGT 2340  
Qy 2363 GGTGAATTTGTTTCCAGCGTACAGTGTCTACATGAGGAAAATCCGGAGTGCACAGAGAA 2422  
Db 2341 GGTGAATTTGTTTCCAGCGTACAGTGTCTACATGAGGAAAATCCGGAGTGCACAGAGAA 2400  
Qy 2423 GAAAGCTCTAAGTCAATGTTTATGTTATCATCGACACAGCACTGAAGGAATTTGAATTTGGAG 2482  
Db 2401 GAAAGCTCTAAGTCAATGTTTATGTTATCATCGACACAGCACTGAAGGAATTTGAATTTGGAG 2460  
Qy 2483 TTGGCCAAACCCAGCGAGCAATGCCCAATTTGTGTGTGAGAAAGACTGCTGTTCAACACTGCA 2542  
Db 2461 TTGGCCAAACCCAGCGAGCAATGCCCAATTTGTGTGTGAGAAAGACTGCTGTTCAACACTGCA 2520  
Qy 2543 AGAGTGTAGTCACTGTTTATATATGATCAGAGATGGCTTTTGGTATCTCTGACAAAGAGATG 2602  
Db 2521 AGAGTGTAGTCACTGTTTATATATGATCAGAGATGGCTTTTGGTATCTCTGACAAAGAGATG 2580  
Qy 2603 AAAGACCATGTCAAGCCGAACTCTTTTTCGATCTCTGTGGCGTAG 2644  
Db 2581 AAAGACCATGTCAAGCCGAACTCTTTTTCGATCTCTGTGGCGTAG 2622

<b>RESULT 4</b> US-10-041-007-34 ; Sequence 34, Application US/10041007 ; Publication No. US20020164736A1 ; GENERAL INFORMATION: ; APPLICANT: Matsuda, Seichi P.T. ; APPLICANT: Schepmann, Hala G. ; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase ; FILE REFERENCE: P02081US1 ; CURRENT APPLICATION NUMBER: US/10/041,007 ; CURRENT FILING DATE: 2002-01-07 ; PRIOR APPLICATION NUMBER: US 60/259,881 ; PRIOR FILING DATE: 2001-01-05 ; NUMBER OF SEQ ID NOS: 41 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 34 ; LENGTH: 2445 ; TYPE: DNA ; ORGANISM: Ginkgo biloba US-10-041-007-34			
Query Match 90.3%; Score 2442; DB 14; Length 2445; Best Local Similarity 100.0%; Pred. No. 0; Matches 2442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	203	TCAGCTGCAGAGACTCGTCCAGATCAGCTTCCACAGGAGGAAAGCTTTGTGTCGAGACTT	362
DB	4	TCAGCTGCAGAGACTCGTCCAGATCAGCTTCCACAGGAGGAAAGCTTTGTGTCGAGACTT	63
QY	263	AATGGGGATTATCATCCAGCTGTCTGGAAAGGACGATTTTCATGACCTCTCTAAACCCCT	322
DB	64	AATGGGGATTATCATCCAGCTGTCTGGAAAGGACGATTTTCATGACCTCTCTAAACCCCT	123
QY	323	AATTCACCGGACATCGAAATCAAGCGTCGATGAGACAAATCAATAAAGATCCAGACA	382
DB	124	AATTCACCGGACATCGAAATCAAGCGTCGATGAGACAAATCAATAAAGATCCAGACA	183
QY	383	TTGGTGAAGGAATCCAGTGCATGTTTTCAGTCCATGGGAGCGGTGAAACGATCCATCT	442
DB	184	TTGGTGAAGGAATCCAGTGCATGTTTTCAGTCCATGGGAGCGGTGAAACGATCCATCT	243
QY	443	GCATATGATACAGCTTGGGTGGCAAGATTCGGTCAATTTGACGGCTCTGGTGACCCCAA	502
DB	244	GCATATGATACAGCTTGGGTGGCAAGATTCGGTCAATTTGACGGCTCTGGTGACCCCAA	303
QY	503	TTTCCCAACCGCTTCAATGGATTCTGAACAAATCAACTGCCAGTGGCTCGTGGGGTGA	562
DB	304	TTTCCCAACCGCTTCAATGGATTCTGAACAAATCAACTGCCAGTGGCTCGTGGGGTGA	363
QY	563	GAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCCTGCCCTCCTCACTCTC	622
DB	364	GAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCCTGCCCTCCTCACTCTC	423
QY	623	AAATATGGAATTAAGGGCGACATTCAGTGCAGAAAGGGTTGAGTTGTGAGAAACAC	682
DB	424	AAATATGGAATTAAGGGCGACATTCAGTGCAGAAAGGGTTGAGTTGTGAGAAACAC	483
QY	683	ATGGAAGAAATGAAGACGAAGCTGCAATCACAGGCGCAAGTGGATTTCGAGGTCTGTTT	742
DB	484	ATGGAAGAAATGAAGACGAAGCTGCAATCACAGGCGCAAGTGGATTTCGAGGTCTGTTT	543
QY	743	CCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATTTCTTATCACTCCCTTTC	802
DB	544	CCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATTTCTTATCACTCCCTTTC	603
QY	803	ATCTCCCAATTCACCAAGCGCCAGAAAGCTTCAAAAGATTTCCCTCAATGTTCTT	862
DB	604	ATCTCCCAATTCACCAAGCGCCAGAAAGCTTCAAAAGATTTCCCTCAATGTTCTT	663
QY	863	CATAACCATCAGACGGGTTGCTCTACTCTCTGAGGGTTTGAAGATGTGGTGAATGG	922
DB	664	CATAACCATCAGACGGGTTGCTCTACTCTCTGAGGGTTTGAAGATGTGGTGAATGG	723
QY	923	CAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTAAAGTCCCTGCATCTACT	982
DB	724	CAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTAAAGTCCCTGCATCTACT	783
QY	983	GCTTGTGTTTCATGCACACTCAAAACAAAGATGCCTTCTCAACTTCGTGCTC	1042
DB	784	GCTTGTGTTTCATGCACACTCAAAACAAAGATGCCTTCTCAACTTCGTGCTC	843
QY	1043	AGCAAAATTTGGCGACTACGTTCCCTTGGCAATCCCACTTGTATTTTGAACGCTCTCG	1102
DB	844	AGCAAAATTTGGCGACTACGTTCCCTTGGCAATCCCACTTGTATTTTGAACGCTCTCG	903
QY	1103	GCTGTGCATACAGTTGAACGCTTGGGAATCGATCCCTATTTCAAGAAAGAAATCAAGAA	1162
DB	904	GCTGTGCATACAGTTGAACGCTTGGGAATCGATCCCTATTTCAAGAAAGAAATCAAGAA	963
QY	1163	TCTCTGGATTACGTTTATAGTACTGGGACCGGAGAGGCGTGGATGGGCAAGATGC	1222
DB	964	TCTCTGGATTACGTTTATAGTACTGGGACCGGAGAGGCGTGGATGGGCAAGATGC	1023
QY	1223	AATCCTATTCTGTGATGTCGATGACACTGCCATGGGTCTTAGAATCCTGAGACTTCATGA	1282
DB	1024	AATCCTATTCTGTGATGTCGATGACACTGCCATGGGTCTTAGAATCCTGAGACTTCATGA	1083
QY	1283	TACAAATGATCTTCAGATGTTCTGGAGAAATTTTCAGAGACGAGAAAGGAGACTTCCTT	1342
DB	1084	TACAAATGATCTTCAGATGTTCTGGAGAAATTTTCAGAGACGAGAAAGGAGACTTCCTT	1143
QY	1343	TTTGGCGGTCAAAACCGCAATTTGGTGTGACCGGATTAATCTTAACTTTATAGATGTTCA	1402
DB	1144	TTTGGCGGTCAAAACCGCAATTTGGTGTGACCGGATTAATCTTAACTTTATAGATGTTCA	1203
QY	1403	GTATGTTTTCCGGGAGAAAGATAATGGAAGAGCTTAAGACCTTCACTCAAAATCATCTC	1462
DB	1204	GTATGTTTTCCGGGAGAAAGATAATGGAAGAGCTTAAGACCTTCACTCAAAATCATCTC	1263
QY	1463	CAAAATGCTTCTTCCAAACAAACCGCATTTGATAGTGGGCTTCAAGAGGATCTTCCT	1522
DB	1264	CAAAATGCTTCTTCCAAACAAACCGCATTTGATAGTGGGCTTCAAGAGGATCTTCCT	1323
QY	1523	CGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAAGTATGCCAAGATTGGAGGCA	1582
DB	1324	CGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAAGTATGCCAAGATTGGAGGCA	1383
QY	1583	AGAAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTAA	1642
DB	1384	AGAAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTAA	1443
QY	1643	ATGCTATATGTGAGCAACCAAAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGGTG	1702
DB	1444	ATGCTATATGTGAGCAACCAAAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGGTG	1503
QY	1703	CAGGCTTACACCAAAAGGAGACTCAACACATTTGTGAGTGGTGGGAGAGAAATCGGATTC	1762
DB	1504	CAGGCTTACACCAAAAGGAGACTCAACACATTTGTGAGTGGTGGGAGAGAAATCGGATTC	1563
QY	1763	AATGATCTTACATTTACCCCGCAGCGCTGTGGAATGTATTTCTCAGTGGCGGTTAGT	1822
DB	1564	AATGATCTTACATTTACCCCGCAGCGCTGTGGAATGTATTTCTCAGTGGCGGTTAGT	1623
QY	1823	ATGTTTGGAGCCAGAAATTCGCTGCTTGTAGAAATTTGGCTATGCCCAAGACTTCTTGC	1882
DB	1624	ATGTTTGGAGCCAGAAATTCGCTGCTTGTAGAAATTTGGCTATGCCCAAGACTTCTTGC	1683
QY	1883	GTTTATTTAGACGATCTTTACGACACCCACGAGATCTCTGGATGATCTTAAATGTTCTCT	1942
DB	1684	GTTTATTTAGACGATCTTTACGACACCCACGAGATCTCTGGATGATCTTAAATGTTCTCT	1743
QY	1943	GAAGCGGTCCGAGATGGGATATCTCTGTCTGATAGCGTTCGGGATTAATCAGTTGAAA	2002
DB	1744	GAAGCGGTCCGAGATGGGATATCTCTGTCTGATAGCGTTCGGGATTAATCAGTTGAAA	1803

QY 2003 GTTTGCTTCTAGGCTGTACAAACACAGTGAATGGATTTGGAAAGATGGACTCAAGGAA 2062  
DB 1804 GTTTGCTTCTAGGCTGTACAAACACAGTGAATGGATTTGGAAAGATGGACTCAAGGAA 1863  
QY 2063 CAAGCGGTGATGTGCTGGGCTATCTTCGAAAAGTATGGGAGGCTTGTCTCGCATCGTAT 2122  
DB 1864 CAAGCGGTGATGTGCTGGGCTATCTTCGAAAAGTATGGGAGGCTTGTCTCGCATCGTAT 1923  
QY 2123 ACCAAAGAGCCGAATGGTTCGGCAGCAAGATGATGTGCGCAGCAATTCACGAAATATGTGGAA 2182  
DB 1924 ACCAAAGAGCCGAATGGTTCGGCAGCAAGATGATGTGCGCAGCAATTCACGAAATATGTGGAA 1983  
QY 2183 AATGCCAAGGTCCATAGACATTTGCGACAGTCTGCTAATCTCAATCTTTTCACTGGA 2242  
DB 1984 AATGCCAAGGTCCATAGACATTTGCGACAGTCTGCTAATCTCAATCTTTTCACTGGA 2043  
QY 2243 GAATTAATCTTCTGATTAATTAATTAACAGCAAGTACAGCTTCGGTCAAAATTTCTGCATCTT 2302  
DB 2044 GAATTAATCTTCTGATTAATTAATTAACAGCAAGTACAGCTTCGGTCAAAATTTCTGCATCTT 2103  
QY 2303 GTGTCTTTGACTGACGACTAATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGT 2362  
DB 2104 GTGTCTTTGACTGACGACTAATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGT 2163  
QY 2363 GGTCAATTTGTTTCCAGGCTACAGTCTACATGATGAGGAAATCCGGAGTGCACAGAGAA 2422  
DB 2164 GGTCAATTTGTTTCCAGGCTACAGTCTACATGATGAGGAAATCCGGAGTGCACAGAGAA 2223  
QY 2423 GAAGCTCTAAGTCAATGTTTATGTTATCATGACAAACGCACTGGAAGGAATGAATTTGGAG 2482  
DB 2224 GAAGCTCTAAGTCAATGTTTATGTTATCATGACAAACGCACTGGAAGGAATGAATTTGGAG 2283  
QY 2483 TTGSCCAACCCAGCGAGCAATGCCCAATTTGTGTGAGAAGACTGCTTTCAACACTGCA 2542  
DB 2284 TTGSCCAACCCAGCGAGCAATGCCCAATTTGTGTGAGAAGACTGCTTTCAACACTGCA 2343  
QY 2543 AGAGTGATGACGCTGTTTATATGATGACAGATGGCTTTGGTATCTCTGACAAAGAGATG 2602  
DB 2344 AGAGTGATGACGCTGTTTATATGATGACAGATGGCTTTGGTATCTCTGACAAAGAGATG 2403  
QY 2603 AAAGACCATGTACCGCAACTCTTTTTCGATCCTGTGGCGTAG 2644  
DB 2404 AAAGACCATGTACCGCAACTCTTTTTCGATCCTGTGGCGTAG 2445

## RESULT 5

US-10-041-007-36  
; Sequence 36, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthese  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 2388  
; TYPE: DNA  
; ORGANISM: Ginkgo biloba  
US-10-041-007-36

Query Match 88.2%; Score 2385; DB 14; Length 2388;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 CTTAATCGGATTAATCATCCAGCTGTCTGGAGGACGATTTTCATCGACTCTCTTAACATCC 319  
DB 4 CTTAATCGGATTAATCATCCAGCTGTCTGGAGGACGATTTTCATCGACTCTCTTAACATCC 63

QY 320 CTTAATTTCCACGCGACATCGAAATCAAGCGTGCATGAGACAATCAATAAAGAAATCCAG 379  
DB 64 CTTAATTTCCACGCGACATCGAAATCAAGCGTGCATGAGACAATCAATAAAGAAATCCAG 123  
QY 380 ACATTTGGTGAAGGAATCCAGTGCATGTTTTCAGTCCATGCGGACGGTGAACCAATCCCA 439  
DB 124 ACATTTGGTGAAGGAATCCAGTGCATGTTTTCAGTCCATGCGGACGGTGAACCAATCCCA 183  
QY 440 TCTGCAATATGATACAGCTTGGGTGGCAAGAAATTCGTCATATTCAGCGCTCTCTGTCACCC 499  
DB 184 TCTGCAATATGATACAGCTTGGGTGGCAAGAAATTCGTCATATTCAGCGCTCTCTGTCACCC 243  
QY 500 CAATTTCCCAAAACGCTTCAATGGAATTCGAAACATCACTGCCAGATGGCTCGTGGGT 559  
DB 244 CAATTTCCCAAAACGCTTCAATGGAATTCGAAACATCACTGCCAGATGGCTCGTGGGT 303  
QY 560 GAGGAGTGCATTTTCTGCGGTATGACAGAGTTTAAACACTCTCGCTCGCTCTCTCACT 619  
DB 304 GAGGAGTGCATTTTCTGCGGTATGACAGAGTTTAAACACTCTCGCTCGCTCTCTCACT 363  
QY 620 CTCAAAATATGGAATTAAGGCGGACATTCAGAGTCAGAAAGGGTTCAGTTTGTGAGAAA 679  
DB 364 CTCAAAATATGGAATTAAGGCGGACATTCAGAGTCAGAAAGGGTTCAGTTTGTGAGAAA 423  
QY 680 CACATGGAAGAAATGAAGGACGAGCTGACAATCAGGCCAAAGTGAATTCGAGGTCTGTG 739  
DB 424 CACATGGAAGAAATGAAGGACGAGCTGACAATCAGGCCAAAGTGAATTCGAGGTCTGTG 483  
QY 740 TTTCTGCTAATGTTAGATGAAGCAAAAGCTTGGGATTCCTTCAATCACTCTCTCACT 799  
DB 484 TTTCTGCTAATGTTAGATGAAGCAAAAGCTTGGGATTCCTTCAATCACTCTCTCACT 543  
QY 800 TTTCTGCTAATGTTAGATGAAGCAAAAGCTTGGGATTCCTTCAATCACTCTCTCACT 859  
DB 544 TTTCTGCTAATGTTAGATGAAGCAAAAGCTTGGGATTCCTTCAATCACTCTCTCACT 603  
QY 860 CTTCAATAACCATCAGACGCGTGTCTCTCTCTCTGAGGGTTCGAAAGATTCGAGGTGGAC 919  
DB 604 CTTCAATAACCATCAGACGCGTGTCTCTCTCTCTCTGAGGGTTCGAAAGATTCGAGGTGGAC 663  
QY 920 TGGCAGAGATCAAAATCTTCAATCAAGAGCGGATCAATTTTAAGTCCCTCGCATCT 979  
DB 664 TGGCAGAGATCAAAATCTTCAATCAAGAGCGGATCAATTTTAAGTCCCTCGCATCT 723  
QY 980 ACTGCTTGTCTTCTATGACACACTCAAAACAAACGATCCCTCACTTTCTCAACTCTGTG 1039  
DB 724 ACTGCTTGTCTTCTATGACACACTCAAAACAAACGATCCCTCACTTTCTCAACTCTGTG 783  
QY 1040 CTGAGCAAAATTTGGCGACTACGTTCTTGTCCATTTACCCACTTGTATTTGAACGCTC 1099  
DB 784 CTGAGCAAAATTTGGCGACTACGTTCTTGTCCATTTACCCACTTGTATTTGAACGCTC 843  
QY 1100 TGGGCTGTCATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAAAGAAATCAAA 1159  
DB 844 TGGGCTGTCATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAAAGAAATCAAA 903  
QY 1160 GAATCTCTGATTTACGTTTATAGTACTGCGACGCCGAAAGAGCGGTGGATGGGCAAGA 1219  
DB 904 GAATCTCTGATTTACGTTTATAGTACTGCGACGCCGAAAGAGCGGTGGATGGGCAAGA 963  
QY 1220 TGCAATCTTATCTGATTCGATGACACTGCCATGGGTCTTGAATCTCTGAGACTTCAAT 1279  
DB 964 TGCAATCTTATCTGATTCGATGACACTGCCATGGGTCTTGAATCTCTGAGACTTCAAT 1023  
QY 1280 GGATACATGATCTTTCAGATGTTTCGGAGAAATTCAGACACGAAAGAGAGACTTCTTT 1339  
DB 1024 GGATACATGATCTTTCAGATGTTTCGGAGAAATTCAGACACGAAAGAGAGACTTCTTT 1083  
QY 1340 TGCTTTGCGGCTCAACGCAAAATTTGGGTGACCGATAATCTTAACTTTTATAGATGTTCA 1399  
DB 1084 TGCTTTGCGGCTCAACGCAAAATTTGGGTGACCGATAATCTTAACTTTTATAGATGTTCA 1143



QY 1400 CAAGTATGTTTCCGGGAGAAAAGTAAATGGAAGCTAGACCTTCACTACAAATCAT 1459  
DB 1144 CAAGTATGTTTCCGGGAGAAAAGTAAATGGAAGCTAGACCTTCACTACAAATCAT 1203  
QY 1460 CTCCTAAATGCTTCTTCCCAAAAACAAACGATTTGATAAGTGGCTGTGCAAGAAGATCTT 1519  
DB 1204 CTCCTAAATGCTTCTTCCCAAAAACAAACGATTTGATAAGTGGCTGTGCAAGAAGATCTT 1263  
QY 1520 CTTGAGAGGTGGAGTATGCTATATAAGTATCCGTGGGATAGAGTATGCCAAGATTGGAG 1579  
DB 1264 CTTGAGAGGTGGAGTATGCTATATAAGTATCCGTGGGATAGAGTATGCCAAGATTGGAG 1323  
QY 1580 GCAAGAAGTTACATAGACGAATTTGGATCAATGATGCTGGCTGGGGAAGACTGTAT 1639  
DB 1324 GCAAGAAGTTACATAGACGAATTTGGATCAATGATGCTGGCTGGGGAAGACTGTAT 1383  
QY 1640 AAGATGCTATATGTGAGCAACGAAAATAATTTGGAGCTGGCCAAATTTGGACTTCAATATG 1699  
DB 1384 AAGATGCTATATGTGAGCAACGAAAATAATTTGGAGCTGGCCAAATTTGGACTTCAATATG 1443  
QY 1700 GTGAGGCTTACACCAAAAGAGACTCAACATTTGCTGCTGGGAGAAATCGGGA 1759  
DB 1444 GTGAGGCTTACACCAAAAGAGACTCAACATTTGCTGCTGGGAGAAATCGGGA 1503  
QY 1760 TTCAATGATCTTACATTTCAACCCGCGAGCGCTGTGGAAATGATTTCTCAGTGGCGGTT 1819  
DB 1504 TTCAATGATCTTACATTTCAACCCGCGAGCGCTGTGGAAATGATTTCTCAGTGGCGGTT 1563  
QY 1820 AGTATGTTTGAGCCAGAAATTCGCTGCTGTGAATTCGCTATGCCAAGACTTCTTGCCTC 1879  
DB 1564 AGTATGTTTGAGCCAGAAATTCGCTGCTGTGAATTCGCTATGCCAAGACTTCTTGCCTC 1623  
QY 1880 GCGTTATTTCTAGACGATCTTTACGACCCAGGATCTCTGATGATCTTAAATTTGTC 1939  
DB 1624 GCGTTATTTCTAGACGATCTTTACGACCCAGGATCTCTGATGATCTTAAATTTGTC 1683  
QY 1940 TCTGAAGCGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGTTCCGGGATATCAAGTTG 1999  
DB 1684 TCTGAAGCGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGTTCCGGGATATCAAGTTG 1743  
QY 2000 AAAGTTTGTCTTAGGCTGTACACACAGTGAATGGATTTGGAAAGATGGACTCAAG 2059  
DB 1744 AAAGTTTGTCTTAGGCTGTACACACAGTGAATGGATTTGGAAAGATGGACTCAAG 1803  
QY 2060 GAACAGGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGAGGCTTCTCGCATCG 2119  
DB 1804 GAACAGGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGAGGCTTCTCGCATCG 1863  
QY 2120 TATACCAAGAAGCCGAATGGTCCGACGAAAGTATGTGCCGACATTCACAGAAATATGTG 2179  
DB 1864 TATACCAAGAAGCCGAATGGTCCGACGAAAGTATGTGCCGACATTCACAGAAATATGTG 1923  
QY 2180 GAAAAATCCAAAGTGTCCATAGCACTTGGACAGTGTCTAACTCAATCTTTTTCAC 2239  
DB 1924 GAAAAATCCAAAGTGTCCATAGCACTTGGACAGTGTCTAACTCAATCTTTTTCAC 1983  
QY 2240 GGAGAAATTTACTTCTGATTTACATTTTACAGCAAGTAGACCTTCCGTTCCAAATTTCTGCAT 2299  
DB 1984 GGAGAAATTTACTTCTGATTTACATTTTACAGCAAGTAGACCTTCCGTTCCAAATTTCTGCAT 2043  
QY 2300 CTTGTGCTTTGACTGACGACTAATCAATGACACCAAGACTTACAGGCGGAGAGAAC 2359  
DB 2044 CTTGTGCTTTGACTGACGACTAATCAATGACACCAAGACTTACAGGCGGAGAGAAC 2103  
QY 2360 CGTGGTGAATTTGGTTTCCAGGCTACGTGTACATGAGGGAATTCGGAGTGCACAGAG 2419  
DB 2104 CGTGGTGAATTTGGTTTCCAGGCTACGTGTACATGAGGGAATTCGGAGTGCACAGAG 2163  
QY 2420 GAAGAAGCTCTAAGTCTATGTTTATGGTATCATCGAACCGCACTGAAGGAATTTGAATGG 2479  
DB 2164 GAAGAAGCTCTAAGTCTATGTTTATGGTATCATCGAACCGCACTGAAGGAATTTGAATGG 2223  
QY 2480 GAGTTGGCCAAACCCAGCGAGCAATGGCCCAATTTGTGTGAGAAAGACTGCTGTTCAACACT 2539

DB 2224 GAGTTGGCCAAACCCAGCGAGCAATGGCCCAATTTGTGTGAGAAAGACTGCTGTTCAACACT 2283  
QY 2540 GCAAGAGTGTAGTACAGCTGTTTATATCTACAGAGATGGCTTTGGTATCTCTGCAAAAGAG 2599  
DB 2284 GCAAGAGTGTAGTACAGCTGTTTATATCTACAGAGATGGCTTTGGTATCTCTGCAAAAGAG 2343  
QY 2600 ATGAAGACCATGTGACGCCGAACCTCTTTTCGATCCTGTGGCGTAG 2644  
DB 2344 ATGAAGACCATGTGACGCCGAACCTCTTTTCGATCCTGTGGCGTAG 2388  
  
RESULT 6  
US-10-041-007-38  
; Sequence 38, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 38  
; LENGTH: 2241  
; TYPE: DNA  
; ORGANISM: Ginkgo biloba  
US-10-041-007-38  
  
Query Match 82.8%; Score 2241; DB 14; Length 2241;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 404 ATGTTTTCAGTCCATCGGCGACGGTGAAGCAATCCATCTGCAATATGATACAGCTTGGGTG 463  
DB 1 ATGTTTTCAGTCCATCGGCGACGGTGAAGCAATCCATCTGCAATATGATACAGCTTGGGTG 60  
QY 464 GCAAGAATTTCCGTCAATTTGACCGCTCTGGTGACCCCCCAATTTCCCAAAAGCTTCAATGG 523  
DB 61 GCAAGAATTTCCGTCAATTTGACCGCTCTGGTGACCCCCCAATTTCCCAAAAGCTTCAATGG 120  
QY 524 ATTCTGAACAATCAATGCTGCGAGATGGCTCGTGGGTGAGAGTGCATTTTCTGGCGTAT 593  
DB 121 ATTCTGAACAATCAATGCTGCGAGATGGCTCGTGGGTGAGAGTGCATTTTCTGGCGTAT 180  
QY 584 GACAGAGTTTAAACACTCTCGCTGCTCTCACTCTCAAAATATGGAATTAAGGCGAC 643  
DB 181 GACAGAGTTTAAACACTCTCGCTGCTCTCACTCTCAAAATATGGAATTAAGGCGAC 240  
QY 644 ATTCAAGTGCAGAAAGGGGTGAGTTTGTGAGAAAACACATGGAAGAAATGAAGGAGAA 703  
DB 241 ATTCAAGTGCAGAAAGGGGTGAGTTTGTGAGAAAACACATGGAAGAAATGAAGGAGAA 300  
QY 704 GCTGACATCAACAGGCAAGTGGATTCGAGTCTGCTTCTGCAATGTTAGATGAAGCA 763  
DB 301 GCTGACATCAACAGGCAAGTGGATTCGAGTCTGCTTCTGCAATGTTAGATGAAGCA 360  
QY 764 AAAAGCTTGGGATTTGGATCTTCTTATACCTCCCTTCTATCTCCCAAAATCCCAAAAG 823  
DB 361 AAAAGCTTGGGATTTGGATCTTCTTATACCTCCCTTCTATCTCCCAAAATCCCAAAAG 420  
QY 824 CGCCAGAAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCAACCATCAGAGCGCGTGG 883  
DB 421 CGCCAGAAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCAACCATCAGAGCGCGTGG 480  
QY 884 CTCTACTCTCTGAGGCTTTGCAAGATGTGCGTGCATGCGCAAGAGATCACAATCTTCAA 943  
DB 481 CTCTACTCTCTGAGGCTTTGCAAGATGTGCGTGCATGCGCAAGAGATCACAATCTTCAA 540



```
QY 944 TCAGAGACGGATCATTTTAAAGCTCCCTGCTACTACTCTGCTGTGCTTTCATGCACACT 1003
Db 541 TCAGAGACGGATCATTTTAAAGCTCCCTGCTACTACTCTGCTGTGCTTTCATGCACACT 600
QY 1004 CAAAACAAACGATGCGCTCCACTTTCTCAACTTCTGCTCTAGCAAAATTTGGCGACTACGTT 1063
Db 601 CAAAACAAACGATGCGCTCCACTTTCTCAACTTCTGCTCTAGCAAAATTTGGCGACTACGTT 660
QY 1064 CTTGCGCAATACCACTTGATCTAATTTGAACGCGCTCTGGGCTGTCGATACAGTTGAACGC 1123
Db 661 CTTGCGCAATACCACTTGATCTAATTTGAACGCGCTCTGGGCTGTCGATACAGTTGAACGC 720
QY 1124 TTGGGAATCGATCGCTAATTTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGG 1183
Db 721 TTGGGAATCGATCGCTAATTTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGG 780
QY 1184 TACTGGGACGCCGAAAGAGCGGTGGGATGGCAAGATGCAATCTAATCTCTGATGTCGAT 1243
Db 781 TACTGGGACGCCGAAAGAGCGGTGGGATGGCAAGATGCAATCTAATCTCTGATGTCGAT 840
QY 1244 GACA CTGCCATGGGTCTTAGAATCTCAGACTTCATCGATACAAATGATCTTCAGATGTT 1303
Db 841 GACA CTGCCATGGGTCTTAGAATCTCAGACTTCATCGATACAAATGATCTTCAGATGTT 900
QY 1304 CTGGAGAAATTCAGAGACGAGAAAGGAGACTTCTTTTGGCTTGGCGTCAAAACGCAATTT 1363
Db 901 CTGGAGAAATTCAGAGACGAGAAAGGAGACTTCTTTTGGCTTGGCGTCAAAACGCAATTT 960
QY 1364 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1423
Db 961 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1020
QY 1424 ATAATGGAAGAAGCTAAGACCTTCACCTACAAATCATCTCAGAAATGCTTTCGCAAAAC 1483
Db 1021 ATAATGGAAGAAGCTAAGACCTTCACCTACAAATCATCTCAGAAATGCTTTCGCAAAAC 1080
QY 1484 AACGCATTTGATAGTGGGCTGTCAAGAGATCTTCTCGAGAGGTGGATATGCTATA 1543
Db 1081 AACGCATTTGATAGTGGGCTGTCAAGAGATCTTCTCGAGAGGTGGATATGCTATA 1140
QY 1544 AAGTATCCGTGGCATAGAATATGCCAAGATTTGAGGCAAGAGTTACATAGACCAATTT 1603
Db 1141 AAGTATCCGTGGCATAGAATATGCCAAGATTTGAGGCAAGAGTTACATAGACCAATTT 1200
QY 1604 GGATCAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTTATATGTGAGCAACGAA 1663
Db 1201 GGATCAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTTATATGTGAGCAACGAA 1260
QY 1664 AAATATTTGGAGCTGGGCAAAATTTGGACTTCAATATGTTGACGGCTTACACCAAAAGGAG 1723
Db 1261 AAATATTTGGAGCTGGGCAAAATTTGGACTTCAATATGTTGACGGCTTACACCAAAAGGAG 1320
QY 1724 ACTCAACACATTTGTCAGCTGTGGAGAGAAATCGGGATTTCAATGATCTTACATTCACCCGC 1783
Db 1321 ACTCAACACATTTGTCAGCTGTGGAGAGAAATCGGGATTTCAATGATCTTACATTCACCCGC 1380
QY 1784 CAGCGGCTGTGGAAATGTAATTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCGCT 1843
Db 1381 CAGCGGCTGTGGAAATGTAATTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCGCT 1440
QY 1844 GCTTGTAGAAATGCTATGCCAAGACTTTCGCTCGAGTATCTTAGAGCATCTTTAC 1903
Db 1441 GCTTGTAGAAATGCTATGCCAAGACTTTCGCTCGAGTATCTTAGAGCATCTTTAC 1500
QY 1904 GACACCAACGGATCTCTGGATGATCTTAAATGTTCTCTGAAGCGTCCGAAGATGGAT 1963
Db 1501 GACACCAACGGATCTCTGGATGATCTTAAATGTTCTCTGAAGCGTCCGAAGATGGAT 1560
QY 1964 ATCTCTGTCTGGATAGCGTTCCGGGATAAATCAGTTGAAAGTTTCTCTTAGGCGCTGTAC 2023
Db 1561 ATCTCTGTCTGGATAGCGTTCCGGGATAAATCAGTTGAAAGTTTCTCTTAGGCGCTGTAC 1620
QY 2024 AACACGTGAATGGAATTTGGAAAGATGGACTCAAGGAAACAAAGCCCGTGTATGCTGGGC 2083
```

```
Db 1621 AACACGTGAATGGAATTTGGAAAGATGGACTCAAGGAAACAAAGCCCGTGTATGCTGGGC 1680
QY 2084 TATCTTCGAAAGATATGGGAGGCTTGTCTCGCATCGTATACCAAGAGCCGAATGCTCG 2143
Db 1681 TATCTTCGAAAGATATGGGAGGCTTGTCTCGCATCGTATACCAAGAGCCGAATGCTCG 1740
QY 2144 GCAGCAAAAGTATGTGGCGACATTTCAACGAATATGTGGAAATATGCCAAAGTGTCCATAGCA 2203
Db 1741 GCAGCAAAAGTATGTGGCGACATTTCAACGAATATGTGGAAATATGCCAAAGTGTCCATAGCA 1800
QY 2204 CTTTGGCAGAGTCTGTACTTAAACTCAATCTTTTTCACCTGGAGAAATTTACTTCTGTATTACATT 2263
Db 1801 CTTTGGCAGAGTCTGTACTTAAACTCAATCTTTTTCACCTGGAGAAATTTACTTCTGTATTACATT 1860
QY 2264 TTACAGCAAGTAGACCTTTCGGTCCAAATTTCTGCATCTTGTGTCTTTGACTGGACGACTA 2323
Db 1861 TTACAGCAAGTAGACCTTTCGGTCCAAATTTCTGCATCTTGTGTCTTTGACTGGACGACTA 1920
QY 2324 ATCAATGACACCAAGACTTTACAGGCGGAGAGAAAACGCTGGTGAATTTGGTTTCCAGCGTA 2383
Db 1921 ATCAATGACACCAAGACTTTACAGGCGGAGAGAAAACGCTGGTGAATTTGGTTTCCAGCGTA 1980
QY 2384 CAGTGTCTACATAGAGGGAATATCCGGAGTGCACAGAGAGAAAGCTCTAAGTCTATGTTTAT 2443
Db 1981 CAGTGTCTACATAGAGGGAATATCCGGAGTGCACAGAGAGAAAGCTCTAAGTCTATGTTTAT 2040
QY 2444 GGTTATCATTCGACAAACGCACTTGAAAGGAATTTGAAATTTGGGAGTTGGCCAAACCCAGCGCAAT 2503
Db 2041 GGTTATCATTCGACAAACGCACTTGAAAGGAATTTGAAATTTGGGAGTTGGCCAAACCCAGCGCAAT 2100
QY 2504 GCCCCATTTGTGTGTGAGAGACATGCTGTTCAACA CATCTGCAAGAGTGTGATGCTGTTTAT 2563
Db 2101 GCCCCATTTGTGTGTGAGAGACATGCTGTTCAACA CATCTGCAAGAGTGTGATGCTGTTTAT 2160
QY 2564 ATCTACAGAGATGCTGTTTGTATCTCTGACAAAGAGATGAAACCATGTGAGCCGAAT 2623
Db 2161 ATCTACAGAGATGCTGTTTGTATCTCTGACAAAGAGATGAAACCATGTGAGCCGAAT 2220
QY 2624 CTTTTCGATCCTGTGGCGTAG 2644
Db 2221 CTTTTCGATCCTGTGGCGTAG 2241
```

## RESULT 7

```
US-09-887-586A-55
; Sequence 55, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A11, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 2861
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2606)
; OTHER INFORMATION: abietadiene synthase
```

US-09-887-586A-55									
Query Match 40.3%; Score 1091.4; DB 9; Length 2861;									
Best Local Similarity 68.4%; Pred. No. 0;									
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;									
Qy	350	GTCCGATGAGACATCAATTAAGAGATCCAGACATTTGGTGAAGGAATCCAGTGCATGTTT	409						
Db	324	GTTCGAGCATCAGACGAGAGCGTATCGAGACATTAATATCCGAGATTAAAGAAATGTTT	383						
Qy	410	CAGTCCATGGGCGCGGTGMAACGAATCCATCTGCATATGATACAGCTTTGGGTGCAAGA	469						
Db	384	AGATGTATGGGTATGGCGAAGCAATCCCTCTGCATATGACCTTGGGTGCAAGG	443						
Qy	470	ATTCGCTCAATTGACGGCTCTGGTGCACCCCAATTTCCCAACGCTTCAATGGATTCTG	529						
Db	444	ATTCGAGCAGTTGATGGCTCTGACACCCCTACCTTTCTCTGAGACGGTTGAATGSAATCTT	503						
Qy	530	ACCAATCACTGCCAGATGGCTCGTGGGTGAGGAGTGCATTTTCTGGCGTATCAGAGA	589						
Db	504	CAAAATCAGTTGAAAGATGGTCTTGGGTGAAGATTTCTACTTTCTGGCATATGACAGA	563						
Qy	590	GTTTTAAACACTCTCGCCCTCCCTCACTCTCAAAATATGGAATTAAGGGCGACATTCAA	649						
Db	564	ATACTGGCTACACTTGCATGATTAATTAACCTTACCTCTGGCGTACTGGGGAGACAA	623						
Qy	650	GTGCAGAAAGGGTTGAGTTTGTGAGAAACACATGGGAGAAATGAGACCAAGCTGAC	709						
Db	624	GTACAGAAAGGTATTGAATTTCTTCAGGACACAAGCTGGAAAGATGGAAGATGAAGCTGAT	683						
Qy	710	AATCAGAGCCAAAGTGGATTTCGAGTCTGTTTCTGCAATGTAGATGAACCAAAAAGC	769						
Db	684	AGTCATAGCCCAAGTGAATTTGAATAGTATTTCTGCAATGCTAAAGGAAGCTAAATC	743						
Qy	770	TTGGGATGGATCTTCCTTATACCTCTCCCTTTCATCTCCCAATCCCAATCCCAAGGCCAG	829						
Db	744	TTAGGCTTGGATCTGCTTACGATTTTGCATTTCTGCAATCAATCAATCAATCAATCAATCA	803						
Qy	830	AAAGAGTTCAAAAGATCCCTCAATGTTCTTCAATCAATCAATCAATCAATCAATCAATCA	889						
Db	804	GCTAAGCTTAAAGAGATCCCACTGATGTTCTATGCTTCTTCAACACGTTATTGTAT	863						
Qy	890	TCTCTGGAGGTTTGAAGATGTGGTGTGCTGGCAAGAGATCAAAATCTTCAATCAAGA	949						
Db	864	TCCTTGAAGGTTTACAAGAAATAGTAGACTGGCAGAAATAATGAACTTCAATCCAAAG	923						
Qy	950	GACGGATCAATTTTAAGTCCCTGCTCATCTACTGTGTGCTTCTATGTCACACTCAAAAC	1009						
Db	924	GATGGATCAATTTCTCAGCTCTCGGCATCTACAGCGGCTGTATTTATGCGTACAGGGAAC	983						
Qy	1010	AAACGATGCTCCACTTTCTCAACTTCGTGCTCAGCAAAATTTGGCGACTACGTTCTCTTG	1069						
Db	984	AAAGTGTCTGGATTTCTTGAATTTGCTTGAAGAAATTCGAAACCATGTGCTGTG	1043						
Qy	1070	CATTACCCACTTGAATCTATTGAAACGCTCTGGGCTGTGATACAGTTGAAACGCTTGGGA	1129						
Db	1044	CACATCCGCTTCAATCTATTGAACGTTTGTGGGCGGTTGATACAGTTGAGCGGCTAGGT	1103						
Qy	1130	ATCGATGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGGTACTGG	1189						
Db	1104	ATCGATGCTTATTTCAAGAGGAGATCAAGAGGATTTGATTTATGTTTACAGCCATTGG	1163						
Qy	1190	GACCGGAAAGAGCGGTGGATGGGCAAGATGCAATCTTATCTCTGATGTGATGACACT	1249						
Db	1164	GA---CGAAAGAGCATTTGATGGGCGAGAGAAATCTCTTCTGATATTGATGATACA	1220						
Qy	1250	GCCATGGCTTTAGATCTTGAGACTTTCATGATACAAATGATCTTTCAGATGTTCTGGAG	1309						
Db	1221	GCCATGGGCTTCCGATCTTTGAGATACATGGATACAAATGATCTCTCAGATGTTTAAAA	1280						
Qy	1310	AAATTCAGAGACGAGAAAGAGACTTTCTTTTGTCTTTTGTCCCGGTCAAAACGCAAAATTTGGTGTG	1369						
Db	1281	ACATTTAGAGATGAGATGGGAGTCTTTTGTCTTTTGTCTTCTTCTTCTTCTTCTTCTTCTTCT	1340						
Qy	1370	ACCGATATCTTAACCTTTATAGATGTTTCAAGTATGTTTTCGGGAGAGAAAGATAATG	1429						
Db	1341	ACAGCATGTTAAAGTCATTCGTTTCAATGTTTCAATGTTTTCGGGAGAAACGATCATG	1400						
Qy	1430	GAAGAAGCTAAGACCTTCACATCAAAATCATCTCCAAATGCTTTGCCAAAAACACGCA	1489						
Db	1401	GAAGAAGCAAAACCTGTACCGAAAGGTATCTGAGGAATGCTCTGGAATAATGTGATGCC	1460						
Qy	1490	TTTGATAGTGGGCTGTCAAGAGATCTTCTCGAGAGGTGGAGTATGCTATTAAGTAT	1549						
Db	1461	TTTGCAAAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGATATGCACTCAATAT	1520						
Qy	1550	CCGTGGCATAGAGTATGCCCAAGATTGGAGGCAAGAGTTTACATAGAGCAATTTGGATCA	1609						
Db	1521	CCCTGGCATAGAGTATGCCCAAGTTTGGAGGCTAGAAGCTATATTTGAAACATATGGCCA	1580						
Qy	1610	AATGATGCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAAATAT	1669						
Db	1581	GATGATGCTGGCTTGGAAATCTGATATATATGATGCCATACATTTCGAAATGAAAGTAT	1640						
Qy	1670	TTGGAGCTGGCAATTCGACTTCAATATGTTGACAGGCTTACACAAAGAGAGACTCAA	1729						
Db	1641	TTAGAACTAGCGAAACTGCACTTCAATAGGTGCACTTATACACCAACAGAGCTTCAA	1700						
Qy	1730	CACATTTGTCAGCTGGTGGAGAAATCGGGATTCAATGATCTTACATTCACCCGCGAGCGG	1789						
Db	1701	GATCTTCGAAGGTGGTGGAAATCATCCGTTTCAAGGATCTGAAATTTCACTCGTGAAGCT	1760						
Qy	1790	CCTGTGGAAATGATTTCTCAGTGGCGGTAGTATGTTTGGCCAGAAATTCGCTGCTGT	1849						
Db	1761	GTGACGGAATATATTTCTCACCGCATCTTTATCTTTGAGCCCGAGTTTCTAAGTGC	1820						
Qy	1850	AGAAATGCTATGCAAGACTTTCTGCTCGCAGTTATTTCTAGACGATCTTTACGACACC	1909						
Db	1821	AGAGAGTTTATACAAATTTTCAATTTTCACTGTTATTTTATAGATGATCTTTATGACGCC	1880						
Qy	1910	CACGATCTCTGGATGATCTTAAATTTGTTCTTGAAGCGGTCCGAAGATGGGATCTCT	1969						
Db	1881	CATGATCTTTAGACGATCTTAAAGTTTTCACAGAAATCAGTCAAAAGATGGGATCTATCA	1940						
Qy	1970	GTGCTGGATAGGTTTCGGGATTAATCAGTTGAAAGTTTGTCTTCTAGGGCTGTACAAACACA	2029						
Db	1941	CTAGTGGG---CCAAATGCCACACAAATGAAATATGTTTGTGGTTTCTACAACTACT	1997						
Qy	2030	GTGAATGGATTTGGAAGATGAGCTCAAGGAAACAGGCCGCTGATGTGCTGGCTATCTT	2089						
Db	1998	TTTAAATGATATAGCAAAAGAGAGCGTGAGAGCGGCGGATGTGCTAGGCTACAT	2057						
Qy	2090	CGAAAGTATGGAGGGCTTGTCTCGCATCGTATACCAAGAACCCGAAATGGTCGGCAGCA	2149						
Db	2058	CAAAATGTTTGAAGTCCAACTTGAAGCTTACACGAAAGAGCAANTGGTCTGAAGCT	2117						
Qy	2150	AAGTATGCTCCGACATTTCAACGAATATGTGGAATATGCCAAAGTGTCCATAGCACTTGG	2209						
Db	2118	AAATATGTGCCATCTCTCAATGAATACATAGAGATGCGAGTGTGTCAATAGCATTTGGGA	2177						
Qy	2210	ACAGTCTGATCAAACTCAATCTTTTCTCAGTGGAGAAATTAATCTTCTGATTTACAG	2269						
Db	2178	ACAGTCTGTTCTCATTTAGTGTCTTTTCTCAGTGGGAGGTTCTTACAGATGAAGTACTCTCC	2237						
Qy	2270	CAAGTACGCTTCGGTCCAAATTTCTGCTCTTGTCTTGTGCTTGTGACCGAGCTAAATCAAT	2329						
Db	2238	AAATTTGATGGGAATCTAGATTTCTTCACTCATGGGCTTAACAGGGCGTTTGGTGAAT	2297						
Qy	2330	GACACCAAGATTTACAGGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGC	2389						
Db	2298	GACACCAAACTTATCAGGCAGAGAGAGGTCAAGGTGAGGTGCTTCTGCCATACAAATGT	2357						
Qy	2390	TACATGAGGGAATCCGGAGTGCACAGAGGAAGAGCTCTAAGTCAATGTTTATGGTATC	2449						
Db	2358	TATATGAAGGAGCCATCTTAAATCTCTTGAAAGAAAGCTCTACACATGTCTATAGTGTCT	2417						

Qy	2450	ATCGACAACGCACTGAAGAGAAATTGANTTTGGGAGTTGGCCCAACCCAGCGAGCAATATGCCCCA	2509
Db	2418	ATGGAAAATGCCCCTCGAAGAGTTTGNATAGGGAGTTTGTGAA-----TAACAAAATATACCG	2471
Qy	2510	TTGTGTGTGTGAGAAGACATGCTGTTCCAAACATGCAAGAGTGTATGCAAGCTGTTTTATATGTATC	2569
Db	2472	GATATTATCAAAAGACACTGGTTTTTGAACATGCAAGAATAATGCAACTCTTTTTATATGCAA	2531
Qy	2570	AGAGATGGCTTTTGGTATCTC---TGACAAAGAGATGAAGAACCATGTGCAGCGCAACTCTTT	2626
Db	2532	GGGGATGTTTGACACTATCACATGATATGGAATTAAGAGCATGTCAAAAATTTGCCTC	2591
Qy	2627	TTTCATCTCTGTGGCTAGCATATA	2648
Db	2592	TTCCAACCAAGTTGCTAGATTA	2613

## RESULT 8

```

RESULT 8
US-09-903-012-55
; Sequence 55, Application US/09903012
; Patent No. US2002094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US2002094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 2861
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2606)
; OTHER INFORMATION: abietadiene synthase
US-09-903-012-55

```

Query Match	40.3%	Score 1091.4;	DB 9;	Length 2861;
Best Local Similarity	68.4%;	Pred. No. 0;		
Matches 1574;	Conservative	0;	Mismatches 713;	Indels 15;
				Gaps 4;

350	Qy	GTCGATGAGACAAATCAATAAAGAAATCCAGACATTTGGTGAAGGAAATCCAGTCGATGCTTTT	409
324	Db	GTTTCAGCATCAGACGAAACGCGTATCGAGACATTAATATCCGAGATTAAGAATAATGTTTT	383
410	Qy	CAGTCCATGGCGACGGTGAAACCAATCCATCTGCATATGATACAGCTTGGGTGGCAAGA	469
384	Db	AGATGTATGGGCTATGGCGAAACGAAATCCCTCTGCATATGACATGCTTGGGTAGCAAGG	443
470	Qy	ATTCGGTCAATTGACGGCTCTGGTGCACCCCAATTTCCCAACGCTTCAATGGAATTCGTG	529
444	Db	ATTCAGCAGTTCGATGGCTCTGCAACCCCTCACCTTCTCGAGACGGTTGAATGGATTCCTT	503
530	Qy	AACAAATCAACTGCCAGATGGCTCTGGGGTGAGGAGTGCAATTTTCTGGGCGTATGACAGA	589
504	Db	CAAAATCAGTTGAAAGATGGGCTCTGGGGTGAGGATTCCTACTCTTTGGCATATGACAGA	563
590	Qy	GTTTTTAAACACTCTCGCGTGCCTCCTCACTCTCAAAATATGGAAATAAGGGCGCAATTCAA	649
564	Db	ATACTGGCTACACTTGCATGATTTATACCTTACCTCTGGCGTATCGGGGACACAA	623

Qy	650	GTGCAGAAAGGGGTTGAGTTTGTGAGAAACAACATCGAAGAAATGAAGGACGAAGCTGAC	709
Db	624	GTA CAGAAAGGTATTTGAAATTTCTTCAGGACACAAGCTGGAAAGATGGAAGATGAAGCTGAT	683
Qy	710	AATCACAGGCCAAGTGGATTTCGAGGTCGTGTTTCTGCAATGTTCTGCAATGTTAGATGAAGCAAAAAGC	769
Db	684	AGTCATAGGCCAAGTGGATTGGAATAGTATTTCTGCAATGCTAAAGGAGCTAAATC	743
Qy	770	TTGGGATTTGGATCTTCCCTTATCACTCCCTTTTCATCTCCCAATCCACCAAAAGCGCCAG	829
Db	744	TTAGGCTTGGATCTGCTCTACGATTTGCCAATCTCTGAAACAATCATCGAAAAGCGGAG	803
Qy	830	AAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCATAAACCATCAGACGGCGTTGCTCTAC	889
Db	804	GCTAAGCTTAAAAAGGATTTCCCACTGATGTTTCTCTATGCCCCCTTCCAAACAACGTTATTGTAT	863
Qy	890	TCTCTGAGGGTTTGCAGATGTTGGTGGACTGCGCAGAGATCACAAATCTTCAATCAAGA	949
Db	864	TCCTTGAAGGTTTACAAGAAATAGTAGACTGGCAGAAATAATGAACATTTCAATCCAAG	923
Qy	950	GACGGATCATTTTAAAGTCCCTGCAATCTACTGCTTGTGTTCTTCATGCAACACTCAAAAC	1009
Db	924	GATGGATCATTTCTCAGCTCTCCGGCAATCTACAGCGCTGATTTCAATGCGTACAGGGAAC	983
Qy	1010	AAACGATGSCCTCCACTTTCTCAACTTGTGTCCTCAGCAAAATTTGGCGACTACGTTCTCTGC	1069
Db	984	AAAAAGTCTTGGATTTCTTGAACCTTTGCTTGAAGAAATTCGGAACCAATGTCCTTGT	1043
Qy	1070	CATTACCCACTTGAATCTATTTTGAACGCTCTGGGCTGTGCAATACAGTTGAAACGCTTGGGA	1129
Db	1044	CACATATCCGCTTGATCTATTTTGAACGTTTGTGGCGGCTTGATACAGTTGACGGCGTAGGT	1103
Qy	1130	ATCGATCGCTATTTCAGAAAGAAATCAAGAATCTCTGATTTACGTTTATAGTACTGG	1189
Db	1104	ATCGATCGCTATTTCAGAAAGGAGATCAAGGAAGCATTTGATTTATGTTTACGCCATTGG	1163
Qy	1190	GAGCCGAAAGAGCGGTGGATGGGCAAGATGCAATCTTATTCCTGATGTCGATGACACT	1249
Db	1164	GA --- CGAAAGAGGCATTGGATGGCGGAGAGAACTCTGTTCTTGATTTGATGATACA	1220
Qy	1250	GCAATGGGCTTTAGAATCTTGAGACTTTCATGGATACAATGTATCTTCAGATGTTCTGGAG	1309
Db	1221	GCCATGGGCTTCGAATCTTTGAGATTTACATGATACAAATGTATCTCTCAGATGTTTAAAA	1280
Qy	1310	AAATTCAGACAGCAGAAAGGAGACTTCTTTGCTTTGTCGGTCAAAAGCAAAATGGTGTC	1369
Db	1281	ACATTTAGATGAGAAATGGGAGTTCTTTTGCTCTTTGGGTCAAAACACAGAGAGATT	1340
Qy	1370	ACCGAATATCTTAACCTTTTATAGATTTCAACAGTATGTTTTCGGGAGAAAAAGATAATG	1429
Db	1341	ACAGACATGTTAAACGTCAATCGTTGTTACATGTTTCATTTTCGGGAGAAACGATCATG	1400
Qy	1430	GAAGAAGCTAAGACCTTCACTACAATATCATCTCCAAAATGCTCTTGCCAAAAACAACGA	1489
Db	1401	GAAGAACAACCTCTGTACCGAAAGGATCTGAGGAATGCTCTGGAATAATGGGATGCC	1460
Qy	1490	TTTTGATGATGGGCTGTCAAGAGGATCTTCTCTGGAGAGGTGAGTATGCTATAAGTAT	1549
Db	1461	TTTGACAAATGGGCTTTTAAAAAGAAATTTTCGGGAGAGGTAGAGTATGCACCTCAAAAT	1520
Qy	1550	CCGTGGCATAGAAGTATGCCAAGATTTGGAGGCAAGAAGTTTACATAGAGCAATTTGGATCA	1609
Db	1521	CCCTGGCATTAAGAGTATGCCAAGTTTGGAGGCTTAGAAGCTTATTTGAAAACCTATGGGCCA	1580
Qy	1610	AATGATGCTCTGGCTGGGGAAGACTGTGTATTAAGATGCTATATGTGAGCAACGAAAAATAT	1669
Db	1581	GATGATGTTGGCTTGGAAAAACCTGTATATATCATGTCATACATTTTCAATGAAAGATAT	1640
Qy	1670	TTGAGCTGCCAAAATTGGACTTCAATATATGGTGCAGGCTTTACACCAAAAGGAGACTCAA	1729
Db	1641	TTAGAACTAGCGAACTGGACTTCAATAAGGTCAGCTTATACACCAACAGAGCTTCAA	1700

QY	1730	CACATTTGTCAGCTGTGGAGAGAAATCGGATTCATGATCTTACATTCACCGCCAGCGG	1789
Db	1701	GATCTTCGAAGGTGGTGGAAATCATCCGGTTTTCAGGATCTGAATTTCACTCGTGAGCGT	1760
QY	1790	CTGTGGAAATGATTTCTCAGTGGCGGTAGTATGTTTGGAGCAGAAATTCGCTGCTGT	1849
Db	1761	GTGACGGAAATATATTTCTCACCGCATCCCTTATCTTTGAGCCGAGTTTCTAAGTGC	1820
QY	1850	AGAAATGGCTATGCCAAGACTTTCTTCCCTCGCAGTTATCTTAGAGCATCTTTACGACACC	1909
Db	1821	AGAGAGGTTTATACAAAACCTTCCAATTTCACTGTTATTTAGATGATCTTTATGACGCC	1880
QY	1910	CACGATCTCTGGATGATCTTAAATGTTCTCTGAGCGGTCCGAGATGGGATATCTCT	1969
Db	1881	CATGGATCTTTAGACGATCTTAAAGTTGTTTACAGAAATCAGTCAAAAGATGGGATCTATCA	1940
QY	1970	GTGCTGGATAGCGTTCCGGATATCAGTTTGAAGTTTGTCTTAGGGCTGTACAAACA	2029
Db	1941	CTAGTGGT---CCAAATGCCACAACAATGAAATATGTTTGTGGGTTTCTACAATACT	1997
QY	2030	GTGAATGGATTTGGAAAGATGAGCTCAGAGAAACAAGCCGATGATGCTGGCTATCTT	2089
Db	1998	TTTAATGATATAGCAAAAGAGAGCGTGAGAGGCAAGGGCGGATGTGCTAGGCTACAT	2057
QY	2090	CGAAAGTATGGAGCGCTTGTCTCGCATCTATACCAAGAGCCGAATGGTCGGCAGCA	2149
Db	2058	CAAAATGTTTGGAAAGTCCAACTTGAGCTTACAGAAAGACAGATGGTCTGAGCT	2117
QY	2150	AAGTATGTGCCGACATTTCAACGAATATGTGSAATATGCCAAATGTCCATAGCACTTGG	2209
Db	2118	AAATATGTGCCATCTTCAATGAATACATAGAGATGCGAGTGTGTCAATAGCATTTGG	2177
QY	2210	ACAGTGTACTAACTCAATCTTTTCACTGGAGATTAATCTTCTGATTTACATTTTACAG	2269
Db	2178	ACAGTGTCTTCAATAGTCTCTTTTCACTGGGAGGTTCTTACAGATGAAGTACTCTCC	2237
QY	2270	CAAGTAGACTTCCGGTCCAAATTTCTGCTATCTTGTCTTGTGACTGGAGCACTAATCAAT	2329
Db	2238	AAATTTGATCGGAACTAGATTTCTTCAACTCATGGCTTAAACAGGGGTTTGGTGAAT	2297
QY	2330	GACACAAAGATTTACAGGCGGAGAGAAACCGTGGTGAATGTTTCCAGCGTACAGTGC	2389
Db	2298	GACACAAAACTTATCAGGCGAGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATACAAATGT	2357
QY	2390	TACATCAGGGAATTCGGAGTGCACAGAGAGAGAGAGCTTAAGTCAATGTTTATGATATC	2449
Db	2358	TATATGAGAGACCATCTTAAATCTCTGAGAGAGAGCTTCAACATGCTATAGTGTCT	2417
QY	2450	ATCGACACGCACTGAAGGAATTTGAATTTGGAGTTGGCCCAACCCAGGAGCAATGCCCA	2509
Db	2418	ATGGAATAATGCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAAAATACCG	2471
QY	2510	TTGTGTGAGAGACTGTGTTTCAACACTGCAAGAGTATGATGAGCTGTTTATATATGATC	2569
Db	2472	GATATTTACAAAAGACTGGTTTGTGAACCTGCAAGAAATATGCACTCTTTTATATGCA	2531
QY	2570	AGAGATGGCTTGTGATCTC---TGACAAAGAGATGAAGACCATGTGACCGCAACTCTT	2626
Db	2532	GGGATGGTTTGACACTATCACATGATATGAAATTAAGAGCATGTCAAAAATTTGCCCTC	2591
QY	2627	TTCCATCCTGTGGGTAGCATTA	2648
Db	2592	TTCCACCAAGTTGCCTAGATTA	2613

RESULT 9

US-09-900-797-55

Sequence 55, Application US/09900797

Publication No. US20030087406A1

GENERAL INFORMATION:

APPLICANT: Chappell, Joseph

APPLICANT: No. US20030087406A1, Joseph P.

APPLICANT: Starks, Courtney M.

; APPLICANT: Manna, Kathleen R.			
; TITLE OF INVENTION: SYNTHASES			
; FILE REFERENCE: 07678-025001			
; CURRENT APPLICATION NUMBER: US/09/900,797			
; CURRENT FILING DATE: 2001-07-06			
; PRIOR APPLICATION NUMBER: US/09/398,395			
; PRIOR FILING DATE: 1999-09-17			
; PRIOR APPLICATION NUMBER: 60/130,628			
; PRIOR FILING DATE: 1999-04-22			
; PRIOR APPLICATION NUMBER: 60/150,262			
; PRIOR FILING DATE: 1999-08-23			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 55			
; LENGTH: 2861			
; TYPE: DNA			
; ORGANISM: Abies grandis			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (3)...(2606)			
; OTHER INFORMATION: abietadiene synthase			
US-09-900-797-55			
Query Match 40.3%; Score 1091.4; DB 10; Length 2861;			
Best Local Similarity 68.4%; Pred. No. 0;			
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;			
QY	350	GTGCGATGAGACAATCAATAAAGAAATCCAGACATTTGGTGAAGGAAATCCAGTGCATGTTT	409
Db	324	GTTCGAGCATCAGACGAGAGCGGTATCGAGACATTAATATCCGAGATTAAAGAAATATGTTT	383
QY	410	CAGTCCATCGGGCGAGCGGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTGGCAAGA	469
Db	384	AGATGTATGGGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTTGGGTAGCAAGG	443
QY	470	ATTCCGTCAATTTGACGGCTCTGTCGACCCCAATTTCCCAACCGTTCCTCAATGGATTCTG	529
Db	444	ATTCCAGCAGTTGATGGCTCTGACAACTCTACTTTCTGAGACGGTTGATGGATTCTT	503
QY	530	AAACAATCAACTCCAGATGGCTCGTGGGTGAGGAGTGCATTTTCTGCGGTATGACAGA	589
Db	504	CAAAATCAGTTGAAAGATGGGTCTTGGGTGAAGGATTTCTACTTCTTGCCATATGACAGA	563
QY	590	GTTTTAAACACTCTCGCTCGCTCTCTCACTCTCAAAATATGAAATGAAGCGGCACATTC	649
Db	564	ATACTGGCTACACTTGCATGTATTTATCCCTTACCTCTGCGGTACTGGGAGACACAA	623
QY	650	GTGCAGAAAGGGGTTGAGTTTGTGAGAAACACATGGAAGAAATGAAGAGCAAGCTGAC	709
Db	624	GTACAGAAAGGTATTGAATTTCTCAGGACACAAGCTGGAAGATGGAAGTGAAGCTGAT	683
QY	710	AATCAGAGCCCAAGTGGATTTCGAGGTGCTGTTTCTGCAATGTTAGATGAAGCAAAAAGC	769
Db	684	AGTCATAGGCCCAAGTGGATTGAAATAGTATTTCTGTCAAATGCTTAAAGAAAGCTTAAATC	743
QY	770	TTGGGATTCGATCTTCTTATCACTCCCTTCTCATCTCCCAATCCACCAAAAGCGCCAG	829
Db	744	TTAGGCTTGGAUTCTGCTTACGTTTCCATTTCCATGCAAAATCATCGAAAAGCGGAG	803
QY	830	AAAAAGCTTCAAAAGATTCCTTCAATGTTTCTTCAATCCATCAGAGCGGCTGCTCTAC	889
Db	804	GCTAAGCTTAAAGGATTTCCCACTGATGTTCTCTATGCCCTTCCAAACACTTATTGTAT	863
QY	890	TCCTCGAGGTTTGGCAAGATGTTGGTGGCTGCAAGAGATCAAAATCTTCAATCAAGA	949
Db	864	TCCTTGGAGGTTTACAGAAATAGTAGACTGGCAGAAATATGAAACTTCAATCCAG	923
QY	950	GACGGATCATTTTAAAGCTCCCTGTCATCTACTGCTTGTGTTGTTCTTATGCACTCAAAAC	1009
Db	924	GATGGATCATTTCTCAGCTCTCCGGCATCTACAGCGCTGTATTCATGCGTACAGGGAAC	983
QY	1010	AAACGATGCCTCCACTTTTCTCAACTTCGCTGCTCAGCAAAATTTGGCGACTACGTTCTTGC	1069

Db 984 AAAAGTCTGGATTTCTTGAATTTTGTCTTGAAGAAATTCGGAACCATGTGCTTGT 1043  
Qy 1070 CATTACCACTGTGATTTGAAAGCCTCTGGGCTGTGCGATACAGTTGAACGCTGGGA 1129  
Db 1044 CACTATCCGCTGTGATCTATTTGAACGTTGTGGCGGTGTGATACAGTTGAGCGCTAGGT 1103  
Qy 1130 ATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGTACTGG 1189  
Db 1104 ATCGATCGCTATTTCAAGAGAGGATCAAGAAAGCATTTGGATTATGTTTACAGCCATGG 1163  
Qy 1190 GACGCCGAAAGAGCGGTGGGATGGGCAAGATGCAATCTATTTCTGTGATGTGATGACACT 1249  
Db 1164 GA---CGAAAGGCAATGGATGGGAGAGAGATCTGTCTCTGATATTGATGATACA 1220  
Qy 1250 GCCATGGCTTTAGATCTCTGAGACTTCATGATACAAATGATCTTCAAGATGTTCTGGAG 1309  
Db 1221 GCCATGGCCCTTCGAATCTTGAGATTACATGGATACAAATGATCTCTCAGATGTTTAAAA 1280  
Qy 1310 AATTTGAGACGAGAAAGGAGACTCTTTTGTCTTTCGGCTCAAGCAAAATTTGGTGTG 1369  
Db 1281 ACATTTAGAGATGAGATGGGAGTCTTTTGTCTTTCGGTCAACACAGAGAGGAGTT 1340  
Qy 1370 ACCGATAATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGGGGAGAAAAGATAATG 1429  
Db 1341 ACAGACATGTTAAAGCTCAATCGTTGTTTCAATGTTTTCGGGGAGAAACGATCATG 1400  
Qy 1430 GAAGAGCTAAGACCTTCACTACAAATCATCTCAAAATGCTTTCGCAAAAACAAAGCA 1489  
Db 1401 GAAGAGCAAAACTCTGTACCGAAAGGTATCTGAGGAATGCTCTGGAATAATGTGGATGCC 1460  
Qy 1490 TTTGATAGTGGCTGTCAAGAGGATCTTCTCGAGAGGTGAGTATGCTATTAAGTAT 1549  
Db 1461 TTTGACAAATGGGCTTTTAAAGAAATATCTCGGGAGAGGTAGAGTATGCACTCAAAATAT 1520  
Qy 1550 CCGTGGCATAGAGTATGCCAAGATTGGAGCAAGAGTTACATAGAGCAATTTGGATCA 1609  
Db 1521 CCCTGGCATAGAGTATGCCAAGGTTGGAGGCTAGAGCTATATTGAACATATGGGCCA 1580  
Qy 1610 AATGATGCTGGCTGGGAAGACTGTGATTAAGATGCTATATGTAGAGCAACGAAAAATAT 1669  
Db 1581 GATGATGTGGCTTGGAAAACTGTATATATGATGCCATACATTTTCAATGAAAAATAT 1640  
Qy 1670 TTGAGCTGGCCAAATTTGACTTCAATATGCTGAGGCTTACACCAAAAGGAGACTCAA 1729  
Db 1641 TTAGAATCTAGGCAAACTGAGCTTCAATAAGGTGAGTCTATACACCAAAAGAGCTTCAA 1700  
Qy 1730 CACATTTGCTGCTGGAGAGAAATCGGGAATCAATGATCTTACATTCACCCGCGAGCGG 1789  
Db 1701 GATCTTGAAGTGGTGGAAATCATCCGGTTTACGGATCTGAATTTTCACTGCTGAGCGT 1760  
Qy 1790 CCTGTGGAATGTATTTCTCAGTGGCGGTAGTATGTTTGAAGCAGAAATTCGCTGCTGT 1849  
Db 1761 GTGACGGAATATATTTCTCACCGGCATCTTTTATCTTTTGAAGCCGAGTTTCTTAAGTGC 1820  
Qy 1850 AGAATTCGCTATGCCAGACTTCTTGGCTCCGAGTATTTCTAGACGATCTTTAGACACC 1909  
Db 1821 AGAGAGGTTTATACAAAACCTTCCAAATTTCACTGTTATTTTATAGATGATCTTTATAGCGCC 1880  
Qy 1910 CACGATCTCTGGATGATCTTAAATTTGTTCTGAAGCGTTCGGAAGATGGATATCTCT 1969  
Db 1881 CATGATCTTTAGACGATCTTAAATTTGTTTACAGATCAGTCAAAAGATGGATCTATCA 1940  
Qy 1970 GTGCTGGATAGCTTGGGATAATCAGTTGAAAGTTTGTCTTCTAGGCGGTGACAAACA 2029  
Db 1941 CTAGTGA---CCAAATGCCAACAAATGAAATATATGTTTGTGGGTTTCTCAAAACT 1997  
Qy 2030 GTGAATGATTTGAAAAGATGGACTCAAGAAACAGGCCGTGATGTCTGGGCTATCTT 2089  
Db 1998 TTTAATGATATAGCAAAAAGAGGACGTGAGAGGCAAGGGCGCGATGTGCTAGGCTACAT 2057  
Qy 2090 CGAAAAGTATGGAGGGCTCTCGCATCGTATACCAAGAGCGGAAATGCTCGCGAGCA 2149  
Db 2058 CAAATGTTTGGAAAGTCAAATCTTGAAGCTTACACGAAAGAGCAAGATGCTCTGAAGCT 2117

Qy 2150 AAGTATGTGCGGACATTTCAAGAAATATGTGGAATAATGCCAAAAGTGTCTATAGACTTGGC 2209  
Db 2118 AAATATGTGCGCATCTTCAATGAATACATAGAGAAATCGAGTGTGTCAATAGCATTTGGGA 2177  
Qy 2210 ACAGTCTGTACTAAACTCAATCTTTTTCACCTGGGAATTTACTTCTCTGATTTACATTTTACAG 2269  
Db 2178 ACAGTCTGTCTCATTTAGTGTCTTTTCTACTGGGAGGTTCTTACAGATGAAGTACTCTCC 2237  
Qy 2270 CAAGTAGACCTTTCGGTCCAAATTTCTGCATCTTGTGTCTTTGACTGGACGACTAATCAAT 2329  
Db 2238 AAAATTGATCGCGAATCTAGATTTTCTTCAACTCATGGCTTAAACAGGCGGTTTGTGTGAAT 2297  
Qy 2330 GACACGAAGCTTACACAGGCGGAGAGAAACCGTGTGAAATTTGGTTTCCAGCGTACAGTGC 2389  
Db 2298 GACACCAAACTTATCAGGCGAGAGAGGTCAAGGTGAGGTGGCTTCTGCGCATCAATGT 2357  
Qy 2390 TACATGAGGGAATAATCCGAGTGCACAGAGAGAAAGCTCTTAAGTCAATGTTTATGTTATC 2449  
Db 2358 TATATGAAGAGCCATCTTAAATCTCTGAAGAGAGACTCTTCAACATGTCTATAGTCTC 2417  
Qy 2450 ATCGACAACGACTGAAGAGAAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATTCGCCA 2509  
Db 2418 ATGGAATAATGCCCTCGAAGAGTTGAATAGGAGTTTCTGTGAA-----TAACAAAAATACCG 2471  
Qy 2510 TTGTGTGTGAGAGACTGTCTGTTTCAACACTGCAAGAGTGTAGCTGTTTATATATGTAC 2569  
Db 2472 GATATTTTACAAAAGACTGGTTTGTGAAACTGCAAGAAATAAATGCAACTCTTTTATATGCA 2531  
Qy 2570 AGAGATGGCTTGTGTATCTC--TGACAAAAGAGATGAAAGACCATGTGACGCGAATCTTT 2626  
Db 2532 GGGATGGTTTGACACTATCACATGATGATGAAATTAAGAGCATGTCAAAAAATTTGCCTC 2591  
Qy 2627 TTGATCTCTGCGGTAGCATATA 2648  
Db 2592 TTCAACACAGTTGCTTAGATTA 2613

## RESULT 10

US-10-041-018-365  
; Sequence 365, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041,018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 365.  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
US-10-041-018-365

Query Match 40.3%; Score 1091.4; DB 12; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

Qy 350 GTGATGAGACAATCAATAAAAGAAATCCAGACATTTGGTGAAGGAAATCCAGTGCATGTTT 409  
Db 324 GTTGACCATCAGACGAGAGCGGTATCGAGACATTAATATCCGAGATTGAAGATATGTTT 383  
Qy 410 CAGTCCATGGCGGACGCGTGAACCGAATCCATCTGTCATATGATACAGCTTGGGTGGCAAGA 469  
Db 384 AGATGTATGGCTATATGCGGAAACCGAATCCCTCTGTCATATGACACTGTCTGGGTAGCAAGG 443  
Qy 470 ATTCCGTCAATTGACGGCTCTGGTGACCCCAATTTTCCCAAGCGCTTCATGATGATCTG 529







QY 1970 GTGCTGGATAGCGTTGGGATATCAGTTGAAAGTTTCTTCTAGGGCTGTACACACA 2029  
DB 1941 CTAGTGGG---CCTAAATGCCCAACAAATGAAATATATGTTTGGGGTTCTTACAAATCT 1997  
QY 2030 GTGAATGATTTGGAAAAGATGGAATCAAGGAAACAAGGCCGTGATGTGCTGGCTATCTT 2089  
DB 1998 TTTAATGATATAGCAAAAGAGGAGCTGAGAGGCAAGGGCGGATGTCTAGGCTACATT 2057  
QY 2090 CGAAAAGTATGGAGGGCTTGTCTGCATCGTATACCAAGAAAGCCGAATGTCGGCAGCA 2149  
DB 2058 CAAATGTTTGGAAAGTCCAATCTGAAGCTTACACGAAAGAGCAAGATGTCGTAAGCT 2117  
QY 2150 AAGTATGTCGACATTTCAACGAATATGTGGAATATGCCAAGTGTCTCATAGCACTTGGC 2209  
DB 2118 AAATATGTCGCATCTCTCAATGATACATAGAGAATGAGTGTCTAATAGCATTTGGGA 2177  
QY 2210 ACAGTCTGTAATAAATCTCAATCTTTTCTACTGAGAAATTAATCTTCTGATTAATTTTACAG 2269  
DB 2178 ACAGTCTGTTCTCATTTAGTGTCTTTTCTACTGGGAGGTCTTTACAGATGAAGTACTCTCC 2237  
QY 2270 CAAGTAGACCTTCGGTCCAAATTTCTGATCTTTGATCTTTGACTGGAGCACTAATCAAT 2329  
DB 2238 AAAATTTGATCGCAATCTAGATTTTCTTCAACTCATGCGGCTTAAACAGGGGTTTGGTGAAT 2297  
QY 2330 GACACCAAGACTTACAGGCCGAGAGAAACCGTGTGAAATTTGGTTTCCAGCGTACAGTGC 2389  
DB 2298 GACACCAAACTTATCAGGCAGAGAGGCTCAAGGTGAGGTGGCTTCTGCCATACAATGT 2357  
QY 2390 TACATGAGGAAATATCCGAGGTGCACAGAGGAAGACTTAAGTCAATGTTTATGGTATC 2449  
DB 2358 TATATGAAGGACCATCTCTAAATCTCTGAAGAAGAGCTCTCAACATGTCTATAGTGTCT 2417  
QY 2450 ATCGCAACGCACTGAAGCAATTTGAATGGGAGTTGGCCAAACCCAGCAGCAATGCCCA 2509  
DB 2418 ATGGAAAATGCCCCGCAAGAGTTGAATAGGAGTTTGTGAA-----TAAACAAATACCG 2471  
QY 2510 TTGTGTGTGAGAGACTGTCTTTTCAACACTGCAAGAGTCAATGAGTGTCTTTTATATGTAC 2569  
DB 2472 GATATTTACAAAGACTGGTTTTTGAATCTGCAAGAAATAATGCAACTCTTTTATATGCAA 2531  
QY 2570 AGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTCTAGCCGAATCTTT 2626  
DB 2532 GGGGATGGTTTGTACATATACATGATATGGAATTAAGAGCATGTCAAAATATGCCCTC 2591  
QY 2627 TTCGATCTCTGTGGCTAGCATATA 2648  
DB 2592 TTCCAACCAAGTTCCTAGATTATA 2613

RESULT 12  
US-10-041-007-3  
; Sequence 3, Application US/10041007  
; Publication NO. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthese  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041.007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
US-10-041-007-3

Query Match 40.3%; Score 1091.4; DB 14; Length 2861;  
Best Local Similarity 68.4%; Pred. No. 0;  
Matches 1574; Conservative 0; Mismatches 713; Indels 15; Gaps 4;

QY 350 GTCGATGAGACAATCAATAAAGAAATCCAGACATTTGGTGAAGAAATCCAGTGTGATGTTT 409  
DB 324 GTTGCACATCAGACGAGAACGGTATCGAGACATTAATATATCCGAGATTAAGAATATGTTT 383  
QY 410 CAGTCCATGGCGAGCGGTGAAAACGAATCCATTCATATATATATATATATATATATATAT 469  
DB 384 AGATGTATGGGCTATGGCGAAACGAATCCCTCTCATATATGACATGCTTTGGGTAGCAAG 443  
QY 470 ATTCGGTCAATTTGACGGCTCTGGTGTGACCCCAATTTCCCAAAACGCTTCAATGGAATCTG 529  
DB 444 ATTCAGAGATTTGATGGCTCTGACAAACCTTCACTTCTCTGAGACGGTTGAATGGAATCTT 503  
QY 530 AACAACTCAACTGCCAGATGGCTCTGGGTGAGAGTGCATTTTCTGGCGTATGACAGA 589  
DB 504 CAAATCAGTTGAAAGATGGGCTTTGGGGTGAAGGATTTCTACTTCTTGGCATATGACAGA 563  
QY 590 GTTTTAAACACTCTCGCTCGCTCTCTCACTCTCAAAATATGMAATAAGGGGAGACATTTCAA 649  
DB 564 ATACTGGCTACACTTTCATGATTTATTAACCTTACCTCTGCGGTACTCTGGGAGACACA 623  
QY 650 GTGCAGAAAGGGTTGAGTTTGTGAGAAACACATGGAAGAAATGAAAGCAAGACTGAC 709  
DB 624 GTACAGAAAGGTATTGAATTTCTTCAAGGACACAAGCTGGAAGATGGAAGATGAAGCTGAT 683  
QY 710 AATCAGAGCCAGTGGATTCGAGTCTGTGTTCTGCAATGTTAGATGAAGCAAAAAGC 769  
DB 684 AGTCATAGGCCAAGTGGATTTGAAATAGTATTTCTGCAATGCTAAAGAGCTAAATC 743  
QY 770 TTGGGATTTGGATTTCTTATCACCCTCTCTTCACTCTCCCAATCCCAAAAGCGCCAG 829  
DB 744 TTAGCTTGGATCTGCTTACGATTTGCAATTTCTTGAACAAATCATCGAAAGCGGGAG 803  
QY 830 AAAAGCTTCAAAAGATTTCCCTCAATGTTTCTTCAATACCATCAGACGGCTTCTCTAC 889  
DB 804 GCTAAGCTTAAAGGATTTCCCACTGATGTTCTATGCCCCCTTCCAAACACGTTATTGAT 863  
QY 890 TCTCTGGAGGTTTCCAGAGTGTGTGACTCGCAAGAGATCAAAATCTTCAATCAAGA 949  
DB 864 TCTTTGGAAGGTTTACAGAAATAGTAGCTGGCAGAAATATGAACTTCAATCCAAG 923  
QY 950 GACGATCAATTTTAAAGCTCCCTCATCTACTGTTGTCTTCTCATGACACTCAAAAC 1009  
DB 924 GATGATCAATTTCTCAGCTCTCCGCACTCAGAGGGCTGTATTCATGCGTACAGGAAAC 983  
QY 1010 AAACGATGCTCCACTTTTCTCAACTTCGTGCTCAGCAAAATTTGGCGACTACGTTCTTGC 1069  
DB 984 AAAAAGTGTCTGGATTTCTTGAACTTTGTCCTTGAAGAAATTCGGAACCATGTGCTTGT 1043  
QY 1070 CATTACCCACTTGTATCTATTTGAACGCTCTGGGCTGTGATACAGTGTGAACGCTTGGGA 1129  
DB 1044 CACTATCCGTTGATCTATTTGAACGTTTGTGGCGGTTGATACAGTGTGAGCGGCTAGGT 1103  
QY 1130 ATCGATCGCTATTTCAAGAAAGAAATCAAGATCTCTGGATTTACGTTTATAGTACTGG 1189  
DB 1104 ATCGATCGCTATTTCAAGAGGAGATCAAGGAAAGCATTTGGATTTATGTTTACAGCATGG 1163  
QY 1190 GACCGCAAGAGCGGTGGGATGGCAAGATGCAATCTTCTTCTGATGTGATGACACT 1249  
DB 1164 GA---CGAAGAGGCAATTTGGATGGCGAGAGAAATCCCTGTTCTGATATTGATGATACA 1220  
QY 1250 GCCATGGGCTTTAGAAATCTCAGACTTTCATGGAATACATGATATCTTCTGAGTGTCTGGAG 1309  
DB 1221 GCCATGGGCTTCAATCTTGAATTTACATGATGATCAATGATATCTCTCAGATGTTTAAAA 1280  
QY 1310 AATTTCAAGACGAGAAAGGAGTCTTTTGTCTTTGGCGGTCAAGCGAAATTTGGTGTG 1369  
DB 1281 ACATTTAGAGATGAGAATGGGAGTCTTTTGTCTTTCTTGGTCTTAAACACAGAGAGGAT 1340  
QY 1370 ACCGATAATCTTAAACGTCATCTGTTGTTTCAATGTTTCAATGTTTCCGGGAGAAACGATCATG 1429  
DB 1341 ACAGACATGTTAAACGTCATCTGTTGTTTCAATGTTTCAATGTTTCCGGGAGAAACGATCATG 1400

Qy 1430 GAAGAAGCTAAGACCTTCACTACAAATCATCTCCAATATGCTCTTGCCAAAACAAAGCA 1489  
Db 1401 GAGAAAGCAAACTCTGTACCGAAGGTATCTGAGAAATGCTCTGGAATAATGTGGATGCC 1460  
Qy 1490 TTTGATAAGTGGCTGTCTCAAGAAGGATCTTCTCGAGAGGTGGAGTAGTATGCTATAAAGTAT 1549  
Db 1461 TTTGACAAATGGCTTTTAAAAGAAATATTCGGGGAGAGGTAGAGTAGTATGCACTCAATAT 1520  
Qy 1550 CCGTGGCATAGAGTATCCCAAGATTGGAGGCAAGAGTTACATAGAGCAATTTGGATCA 1609  
Db 1521 CCCTGGCATAGAGTATCCCAAGGTGGAGGCTAGAAGCTATATTGAAAACCTATGGGCCA 1580  
Qy 1610 AATGATGCTGCTGGGGAAGACTCTGTATTAAGATGCTATATGTGAGCAACGAAAAATAT 1669  
Db 1581 GATGATGTGCTTGGAAAACCTGTATATATGATGCCATACATTTGAAATGAAGAATAT 1640  
Qy 1670 TTGGAGCTGGCCAAATTGGACTTCAATATGTTGAGGCTTTACACCAAAAGAGACTCAA 1729  
Db 1641 TTAGAACTAGCGAACTGCACTTCAATAAGGTGCAGTCTATACACCAAAAGAGCTTCAA 1700  
Qy 1730 CACATGTGCTGGTGGAGAGATCGGGATTCATATGATCTTACATTCACCCGCGAGCGG 1789  
Db 1701 GATCTTCGAAGGTGGTGGAAATCATCCGGTTTCAAGGATCTGAAATTTCACTCGTGAGCGT 1760  
Qy 1790 CCTGTGGAATGTATTTCTCAGTGGCGGTAGTATGTTTGAGCCAGAAATTCGCTGCTGT 1849  
Db 1761 GTGAGGAATATATTTCTCAAGGCACTCTTTATCTTTGAGCCGAGTTTCTTAAGTGC 1820  
Qy 1850 AGAATTCGCTATGCCAAGACTTCTTGCCTCGCAGTTATTTAGAGCATCTTTACGACACC 1909  
Db 1821 AGAGAGTTTATACAAAACCTTCCAAATTTCACTGTATTTAGATGATCTTTATGAGCGC 1880  
Qy 1910 CACGATCTCTGGATGATCTTAAATGTTCTGTAAGCGGTCCGAAGATGGGATATCTCT 1969  
Db 1881 CATGATCTTTAGACGATCTTAAAGTTGTTACAGAAATCAGTCAAAAGATGGATCTATCA 1940  
Qy 1970 GTGCTGGATAGCTTGGGATTAATCAGTTGAAATTTGCTTCTAGGCTGTACACACA 2029  
Db 1941 CTAGTGGG---CCAAATCCCAACAAATAATATATGTTTGTGGGTTTCTACAATACT 1997  
Qy 2030 GTGAATGATTTGAAAAGATGGACTCAAGGAAACAAAGCGGTGATGTCTGGGCTATCTT 2089  
Db 1998 TTTAATGATATAGCAAAAGAGGAGCTGAGAGGCAAGGGCGGATGTGCTAGGCTACAT 2057  
Qy 2090 CGAAAGTATGGAGGGTCTGCTCGCATGTTATACAAAGAGCGAATGTCGGCAGCA 2149  
Db 2058 CAAAATGTTGGAAAGTCAAATCTGAAGCTTTACAGAAAGAGCAGATGCTCTGAAGCT 2117  
Qy 2150 AAGTATGTCGCAATTCACGATATGTGGAATGCAAGTGTCATAGCACTTGGC 2209  
Db 2118 AAATATGTGCCATCTTCAATGAATACATAGAGATGCGAGTGTCTAATAGCATGGGA 2177  
Qy 2210 ACAGTCGTACTAACTCAATCTTTTCACTGGAGAAATTTACTTCTGTGATTAATTTACAG 2269  
Db 2178 ACAGTCGTCTCATTTAGTCTCTTTTCACTGGGAGGTTCTTACAGATGAAGTACTCTCC 2237  
Qy 2270 CAAGTAGACCTTCGGTCCAAATTTCTGCATCTTGTCTTTGACTGGACGACTAATCAAT 2329  
Db 2238 AAAATGTATCGGAATCTAGATTTCTTCAATCTGAGGCTTAAACAGGCGGTTTGTGTGAT 2297  
Qy 2330 GACACCAAGACTTACAGCGCCGAGAGAAACCGTGTGTGAATTTGGTTTTCAGCGGTACAGTGC 2389  
Db 2298 GACACCAAACTTTATCAGCGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATCAATGT 2357  
Qy 2390 TACATGAGGGAATCCCGAGTGCACAGAGGAAGCTCTAAGTCAATGTTTATGGTATC 2449  
Db 2358 TATATGAGGACCATCTTAAATCTTCTGAAGAAGAGCTCTACAACTGTCTTATAGTGC 2417  
Qy 2450 ATCGAACAGCACTGAAGAAATTTGAATTTGGAGTTTGGCCAAACCCAGCGAGCAATGCCCA 2509  
Db 2418 ATGGAATATGCCCTCGAGAGTTGAATAGGAGTTTGTGAA-----TAAACAAATACCG 2471  
Qy 2510 TTGTGTGAGAGAGCTGCTGTTCAACACTGCAAGAGTGTGAGCTGTTTATATGTAC 2569

Db 2472 GATATTTACAAAGACTGGTTTTTGAACCTGCAAGAATAATGCAACTCTTTTATATGCA 2531  
Qy 2570 AGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAAGACCAATGTGACGCCAATCTT 2626  
Db 2532 GGGGATGGTTTGACACTATCACATGATGAAATTAAGAGCATGTCAAATAATTGCTC 2591  
Qy 2627 TTCGATCCTGTGGGTAGCAT 2648  
Db 2592 TTCCAACCACTGTGCTAGATTA 2613  
RESULT 13  
US-09-887-586A-45  
; Sequence 45, Application US/09887586A  
; Patent NO. US20020094556A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20020094556A11, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/887,586A  
; CURRENT FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(2347)  
; OTHER INFORMATION: E-alpha-bisabolene synthase  
US-09-887-586A-45  
Query Match 27.0%; Score 730.2; DB 9; Length 2424;  
Best Local Similarity 59.6%; Pred. No. 3.1e-208;  
Matches 1370; Conservative 0; Mismatches 893; Indels 34; Gaps 7;  
Qy 434 AATCCATCTCATATGATACAGCTTGGGTGSCAAGAAATTCGTCAAATTTGACGGCTCTGGT 493  
Db 134 ACTCCATCTCTTATGACACAGCATGGGTAGCGGGTGCCGCCATTTGATGGCTCTGCT 193  
Qy 494 GCACCCCAATTTCCCAACCGCTTCAATGGATTTCTGAACAAATCAACTGCCAGATGGCTCG 553  
Db 194 CGCCCGCAATTTCCCAACACAGTTGACTGGATTTTGAANAACCAAGTTAAAGATGGTTCA 253  
Qy 554 TGGGTGAGGAGTGCAATTTTTCGGCGTATGACAGAGTTTAAACACTCTCGCCCTCGCTC 613  
Db 254 TGGGGAATTCAGTCCCACTTTTCTGCTCGGACCGTCTTCTTGCCACTCTTCTTGTGTT 313  
Qy 614 CTCACTCTCAAAATATGGAATAGGGCGACATTCAGTGCAGAAAGGGTTGAGTTTGTG 673  
Db 314 CTGTGCTCTTAAATGGAACGTTGGGGAATCTCAAGTAGAGCAGGGAATTTGAATTCATA 373  
Qy 674 AGAAAAACATGGAAGAAATGAAGGACGAAGCTGACAAATCACAG---GCCAAGTGGATTC 730  
Db 374 AAGAGCAATCTGGAACTAGTAAAGGATGAACCGATCAAGTAGCTTGGTAACAGACTTT 433  
Qy 731 GAGTCTGTGTTCTGCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGATCTTCTTAT 790  
Db 434 GAGATCATATTTCTTCTCTGTTAAGAGAAAGCTCAATCTCTCGCCTCGGACTTCCCTAC 493  
Qy 791 CACTCTCTTCTATCTCTCCCAATCCCAAAAGGCCGAGAAAGCTTCAAAAGATTTCC 850







Db	1271	AGAGTGGAGCAAGAGATACTGCTCAAGTGTATCGCCAGATTATGCACGATAGCAAAA	1330
Qy	1631	ACTGTGTATAAGATGCTATATGTGAGCAACGAAAAATATTTGAGCTGGCCAAATTGGAC	1690
Db	1331	TGGCTTTACAAAGCTACCTACGTGGAACAAATGAAAGTTTTTAGAGCTGGGAAATTAGAT	1390
Qy	1691	TTCAATATGGTGCAGGCTTACACCAAGAGGAGCTCAACACATTTCTCAGCTGGTGGAGA	1750
Db	1391	TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAAGAATGTTACCAAGCTGGTTTGA	1450
Qy	1751	GAAATCGGATTTCAATGATCTTACATTCACCCGCGAGCGCTGTGGAATGTTATTTCTCA	1810
Db	1451	GATTCGGGGTGGCACTATTACCTTCGCTCGGAGAGGCGCTGGAACTCTACTTCTTA	1510
Qy	1811	GTGGCGGTAGTATGTTGAGCCAGAAATTCGCTGTGTGTAGAAATTCCTATGCCAAGACT	1870
Db	1511	GTAGCGCGGGACCTATGAACCCAGTATGCCAAATGCCAGTTCTCTTTACAAAAGTG	1570
Qy	1871	TCTTGCTCGCAGTTATCTAGACGATCTTTACGACCCAGGATCTCTGGATGATCTT	1930
Db	1571	GCATGCTTGAGACTGTCTGAGACGATATGTATGACACTTTATGAAACCCCTAGATGAATTG	1630
Qy	1931	AAATTCCTCTCAAGCGGTCCGAAGATGGGATATCTCTGCTGGATAGCGTTTCGGAT	1990
Db	1631	AGCTATTCTCTGAGGCTGTGAGAGATGGGAACTCTCTTTACAGAAACCTTCCAGAC	1690
Qy	1991	AATCAGTTGAAAGTTTCTCTAGGGCTGTACAACACAGTGAATGGAATTTGAAAAGAT	2050
Db	1691	TAT---ATGAACTATGTTTACCAAACTATTTATGACATAGTTCCAGAGGTGGCTTGGAG	1747
Qy	2051	GGACTCAAGGAACAAGCCGTGATGTCTGGGCTATCTTCGAAAAGTATGGGAGGGCTTG	2110
Db	1748	GCAGAGAAGGAACAGGGGCGTGAATTTGGTCAGCTTTTTCAGAAAGGGATGGGAGGATTAT	1807
Qy	2111	CTCGCATCGTATACCAAGAACCGAATGTGCGCAGCAAGTATGTGCCGACATTCACAC	2170
Db	1808	CTTCTGGGTTATTTAGAGAGCTGAATGGTTAGTGTCTGAGTATGTGCTTACCTTTGGAC	1867
Qy	2171	GAATATGTGAAAATGCCAAAGTGTCCATAGCACTTCCGACAGTCTGTAATACTCAATC	2230
Db	1868	GAGTACATAAAGAAATGGAATCACATCTATCGGCCAAGTATCTTCTGTTGAGTGGAGTG	1927
Qy	2231	TTTTTCA---CTGGAGAAATTACTTCTCTGATTTATTTTACAGCAAGTAGACCTTCGGTCC	2287
Db	1928	TTGATAATGATGGGCAACTCCTTTCCGAAGAGGCAATTAGAGAAAGTAGATTATCCAGGA	1987
Qy	2288	AAATTTCTGC-----ATCTGTGTCTTTGACTGGAGCACTAATCAATGACACCAAGACT	2341
Db	1988	AGAGTGTCTCACAGAGCTGAATAGCCTCATTTCCGCCCTGGCGGATGACACGAAGACA	2047
Qy	2342	TACAGCGCGAGAGAAACCGTGTGAATTTGGTTTCCAGCGTACAGTCTACATGAGGGAA	2401
Db	2048	TATAAAGCTGAGAAGGCTCGTGAGAAATTGGCGTCCAGCAATTGAATGTTACATGAAAGAC	2107
Qy	2402	AATCCGGAGTGACAGAGGAAGAGCTCTAAGTCAATGTTTATGGTATCATCGACAAGCA	2461
Db	2108	CATCTCTGAATGTACAGAGGAAGAGGCTCTCGATCACATCTATAGCAATTTCTGGAGCCGGCG	2167
Qy	2462	CTGAAGGAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA	2521
Db	2168	GTGAAGGAATGACAGAGAGTTTCTGAGCCCG---ACGACGTCCCATTTCCGCTCGCAAG	2224
Qy	2522	AGACTGCTGTTCAACACTGCAAGAGTGAATCAGCTGTTTTTATATGTACAGAGATGGCTTT	2581
Db	2225	AAGATGCTTTTCGAGGAGACAAGGTGACCGATGGTGATATTCAGGATGGAGATGGATTC	2284
Qy	2582	GGTATCTCTGACAAAGAGATGAAGACCATGTGACGGGAACTCTTTTCGATCTCTGTGGCG	2641
Db	2285	GGTGTTCCTCAATTTAGAAATCAAAGATCATATCAAAAGAGTGTCTCATTTGAACCGCTGCCA	2344
Qy	2642	TAGCA-----TACTGATATATATATATATTTATATTTCAATCCAAAAAAA	2688
Db	2345	CTGTAATCAAAATAGTTGGCAATAATAATTGAAATAATGTCAACTATATGTTTTCACAAAAA	2404

Qy	2689	AAAAAAAAAAAAAAAAAAAA	2705
Db	2405	AAAAAAAAAAAAAAAAAAAA	2421

Search completed: August 25, 2004, 02:26:20  
Job time : 1187.41 secs

**THIS PAGE IS BLANK**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 06:59:45 ; Search time 6846.72 Seconds  
(without alignments)  
11797.935 Million cell updates/sec

Title: US-10-041-007-1  
Perfect score: 2705  
Sequence: 1 attattgcctaacctgcac.....aaaaaaaaaaaaaaaaaaaaa 2705

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: em\_estba.\*  
2: em\_estum.\*  
3: em\_estin.\*  
4: em\_estu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hcc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hcc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_fod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	380.8	14.1	771	14	CF478561
2	307.6	11.4	730	13	BX682869
3	267.4	9.9	641	14	CF475347
4	264.2	9.8	2797	11	AY109636

5	231.4	8.5	476	10	BP517045
6	228.8	8.5	725	14	CF397293
7	214.8	7.9	583	13	BX681656
8	205.4	7.6	634	14	CF397946
9	204.4	7.6	725	14	CF470471
10	190.8	7.1	700	14	CF474786
11	190.4	7.0	745	14	CF477103
12	188.8	7.0	696	14	CF401916
13	186	6.9	581	13	BX677599
14	184.4	6.8	624	13	BU971902
15	184.2	6.8	688	14	CA223665
16	176.2	6.5	569	10	AW287754
17	175.8	6.5	601	9	AW043070
18	175.2	6.5	833	14	CB671661
19	174.8	6.5	638	14	CA203354
20	174.2	6.4	873	12	BG368010
21	172.8	6.4	669	14	CF479802
22	172.2	6.4	863	14	CB630278
23	171.8	6.4	706	14	CA917162
24	171	6.3	533	13	BQ698297
25	170.4	6.3	616	14	CF663845
26	169.8	6.3	740	14	CF477562
27	169	6.2	576	9	AU299985
28	168	6.2	865	14	CB671672
29	166.2	6.1	804	14	CF666338
30	165.6	6.1	506	12	BG317597
31	164.4	6.1	697	13	BQ196773
32	162	6.0	669	14	CF397861
33	162	6.0	682	14	CF397245
34	158.8	5.9	554	13	BU101395
35	156.6	5.8	599	13	BX680641
36	153.6	5.7	539	10	AW287756
37	152.2	5.6	741	14	CF389118
38	148.2	5.5	591	14	CA305371
39	146.4	5.4	840	14	CB624295
40	145.8	5.4	651	14	CF479884
41	143	5.3	839	14	CB619396
42	143	5.3	844	14	CB619357
43	140.8	5.2	599	12	BG526917
44	140.4	5.2	385	10	AW697531
45	135.4	5.0	843	14	CF208265

## ALIGNMENTS

RESULT 1  
CF478561

LOCUS RTW3\_20\_F04.g1\_A022 Well-watered loblolly pine roots WW3 Pinus taeda CDNA clone RTW3\_20\_F04\_A022 5', mRNA sequence.

DEFINITION CF478561

ACCESSION CF478561.1 GI:34507430

VERSION EST.

KEYWORDS Pinus taeda (loblolly pine)

SOURCE Pinus taeda

ORGANISM Pinus taeda

REFERENCE 1 (bases 1 to 771)

AUTHORS Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gabremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.

TITLE An EST database from well-watered loblolly pine (Pinus taeda) roots

JOURNAL Unpublished (2003)

COMMENT Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the

CF478561 771 bp mRNA linear EST 08-SEP-2003  
RTW3\_20\_F04.g1\_A022 Well-watered loblolly pine roots WW3 Pinus taeda CDNA clone RTW3\_20\_F04\_A022 5', mRNA sequence.

CF478561  
Pinus taeda (loblolly pine)  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus taeda (loblolly pine)  
1 (bases 1 to 771)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gabremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.  
An EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the

BP517045 NXSI\_008  
CF397293 RTDS3\_2\_A  
BX681656 BX681656  
CF397946 RTDS3\_23  
CF470471 RTDS1\_17  
CF474786 RTW2\_7\_B  
CF477103 RTW2\_5\_A  
CF401916 RTW1\_15  
BX677599 BX677599  
BU971902 HB20A06r  
CA223665 SCJFFL1C0  
AW287754 EST0002 S  
AW043070 ST28H10\_P  
CB671661 OSJNE805F  
CA203354 SCSBFL103  
BG368010 HVSME1001  
CF479802 RTW3\_12  
CB630278 OSTIEB07C  
CA917162 EST641309  
BQ698297 NXPV\_068  
CF663845 RTCNT1\_5  
CF477562 RTW3\_8\_G  
AU299985 AU299985  
CB671672 OSJNE805P  
CF666338 RTCNT1\_22  
BG317597 NXPV\_003  
BQ196773 NXLI05\_B  
CF397861 RTDS3\_23  
CF397245 RTDS3\_2\_A  
BU101395 946153G07  
BX680641 BX680641  
AW287756 EST0004 S  
CF389118 RTDR2\_13  
CA305371 haasp004xK  
CB624295 OSTIEB12M  
CF479884 RTW3\_12  
CB619396 OSTIEB03O  
CB619357 OSTIEB03N  
BG526917 NXPV\_057  
AW697531 ST64A12\_P  
CF208265 CAB20003\_

University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES

Location/Qualifiers

1..771  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3152"  
/clone="RTW3.20.F04.A022"  
/lab\_host="pH108-T1 phage-resistant E. coli"  
/clone\_lib="Well-watered loblolly pine roots W3"  
/note="Vector: pSul180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSul180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 14.1%; Score 380.8; DB 14; Length 771;  
Best Local Similarity 68.5%; Pred. No. 8e-57;  
Matches 526; Conservative 0; Mismatches 242; Indels 0; Gaps 0;

375 TCAGACATGGTGAAGAAATCCAGTCATGTTTCAGTCCATTCAGTGGCGAGCGTGAACGA 434  
4 TCAGACATTAATATCTGTTATTAGAGATGTTTCAGAGGTATGGGTGATGGGGAACAA 63  
435 ATCCATCTGCATATGATACAGCTTGGTGGCAGAAATCCGTCAATTCAGCGCTCTGGTG 494  
64 ATCCCTCTGGTATGACATCTGTTGGTAGCAAGATTCAGACATTCGATGGCTGACC 123  
495 CACCCCAATTTCCCAACCGCTCAATGATTTCTGAACATCAATCAATGCGAGATGGCTGT 554  
124 ACCCTCAGTTTCTCAACACTCCATGATGTTCTTCAGAAATCAGCTGCAAGATGGCTCTT 183  
555 GGGGTGAGGATGCAATTTTCTGGCTATGACAGATTTTAAACATCTCGCTGCTCC 614  
184 GGGGTGAAGACAACTATTTCTTAAGATATGACAGATCTCATCTCTTGGTGTATTA 243  
615 TCATCTCTCAAAATATGAATGAAGGGGACATTCAGTGCAGAAAGGGTGTAGTTTGA 674  
244 TTACCTCTACCGTATGGCCGATGCGCAGATCAAGTGCAGAAAGGTCTTGAATCTTCA 303  
675 GAAACACATGGAAGAAATGAAGACGAGCTGACATCAATCAGCCCAAGTGGATTCGAGG 734  
304 AGAAACACGCTGAATGATGGAGTGAAGCGGACATCAACCGGCAAGTGGATTCGAA 363  
735 TCCTGTTTCTCTGCAATGTATAGATGAAGCAAAAGCTTGGATTTGGATCTTCTTATCACC 794  
364 TTGCTTTCTCTGATGCTAAACGAGCAAACTATGCTTTGGATCTGCTTATGAT 423  
795 TCCCTTTTCACTCTCCAAATCCACAAAGCGCCAGAAAGGTTCAAAAGATTCCTCACT 854  
424 TGCAATTCATTAACAACTCATTAAGAAAGGGGAAGCTTAAGCTTAAGAGGATTTCACTG 483  
855 ATGTTCTTCAATACCATCAGCGGCTGTCTCTACTCTCTGAGGGTTCGCAAGATGTGG 914  
484 ATACTCTCTATAGGTTTCCAACTAATATCTTGTATTAATTTGGAGGCTGCAAGATAG 543  
915 TGGACTGGCAAGAGATCACAATCTCAATCAAGAGACGGATCAATTTTAAAGTTCCTCTG 974  
544 TGGAGTGGCCCAAAATCATCAAACTTCAATCCAAAGGATGATCGTTCTCGGCTGGCAG 603  
975 CATCTACTGTTGCTTTCATGACACTCAAAACAAACGATGCTCACTTCTCACT 1034  
604 CATCTACAGCGCTGTATTCATGCGTACAGGAACACCAAGTGTGGAAATCTTGAAT 663  
1035 TCGTGCTCAGCAAAATTTGGGCGACTACGTTCTTGGCATTACCCACTTGTATTTGAAC 1094

664 TGGTCTTAATGAATTCGAAACACAGTCGCGTGTCACTATCCGCTTAATCTATTGAAC 723  
1095 GCCTCTGGCTGTGATACAGTGAACGCTTGGGAATCGATCGCTATT 1142  
724 GTCGTGGGCGCTGTGATACTGTGAAACGCTTAGGATCGATCGCCATT 771

RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..730  
/organism="Pinus pinaster"  
/mol\_type="mRNA"  
/strain="ecotype: Corsican"  
/db\_xref="taxon:71647"  
/clone="122H09"  
/tissue\_type="differentiating xylem"  
/dev\_stage="adult"  
/clone\_lib="Pinus pinaster differentiating xylem adult"  
/note="Vector: Uni-Zap XR lambda (Stratagene); Site 1: Eco RI; Site 2: Xho I; A composite cDNA library was made with mRNA isolated from normal, compression, opposite, early and late wood of Maritime pine uni-directionally cloned into Uni-Zap XR using the ZAP-cDNA Synthesis kit (Stratagene). pBluescript SK(-) plasmids were obtained by in vivo mass excision. The nucleotide sequence of the 5' end was obtained by automated sequencing with the T3 primer by GENOME EXPRESS, Meylan, France"

ORIGIN

Query Match 11.4%; Score 307.6; DB 13; Length 730;  
Best Local Similarity 64.8%; Pred. No. 5.2e-44;  
Matches 473; Conservative 0; Mismatches 254; Indels 3; Gaps 1;

1734 TTGTGAGTGTGGAGAGATCGGATTCATATCATTCACCCGCGAGCGGCTG 1793  
1 TTGGAGGTGGTGGAAATTCATCGGGTTCTCAGAGCTGGGCTTCATCTCGGAGCTGTGG 60  
1794 TGGAAATGTAATTTCTCAGTGGCGGTTAGTATGTTGAGCCAGAAATTCGCTGCTTGTAGAA 1853  
61 CTGAATATATTTTTCACGAGCGTCATTTTGTGAGCCAGAGTTCGCTACCTCGCGAG 120  
1854 TTGCCTATGCCAAGACTTCTTGCTCGCAGTATCTAGACGATCTTTAGCAGACCCACG 1913  
121 CCGTTTACACAAAACCTCTTAATTTTCCAGGTCATATAGATGATCTTTATGACCTCATG 180  
1914 GATCTCTGGATGATCTTAATTTGTTCTCTCAAGCGGTCCGAAAGATGGGATATCTCTGTGC 1973  
181 GAACCTTGGCAATCTCAAGTTGTTTTCAGAACTCTGGAAGATGGGATCTATCTCTTGG 240

```
QY 1974 TGGATAGCGTTCCGGGATATCACTGTTGAAAGTTTCTTCTAGGCTGTACACACAGTGA 2033
D 241 TAGATCAAAAT---GCCACAAGACATGAATAATATGTTTAAAGGGTTTATCAATACTTTTA 297
QY 2034 ATGATATTGGAAAAGATGAGCTCAAGGAACAAGGCCGTGATGCTCGGCTATCTTCGAA 2093
D 298 ATGNAATCGCTGAAGAGGGGGGAAAGCGAGGGCGTGATGCTAGGTGTACATTCAAA 357
QY 2094 AAGTATCGGAGGGCTTCTCGCATCGTATATACAAAGAACCGAATGCTCGGCGACAAAGT 2153
D 358 AAGTTTGGGAGGTCAGCTAGGAGCGTATACCAAGAAAGCAGAATGCTCTGCAGCTAGAT 417
QY 2154 ATGTCGCGCATTCCAAGATATGTGGAATATGGAATGCCAAGTGTTCATAGCATTGCGACAG 2213
D 418 ACGTGCCTCTATGATGAATATATAGGGAACGCGAGTGTTCATATAGCATTTGGGAACAG 477
QY 2214 TCGTACTAAACTCAATCTTTTCACTGGAGAAATPACTTCTCTGATTACATTTTACAGCAAG 2273
D 478 TGGTCTCATCAGCGCTCTTTTACCGGGGAGATCTTACAGATGATATCTCTCCAAA 537
QY 2274 TAGACCTTCGGTCCAAATTTCTGCATCTTGTCTTTGACTGGACGACTAATCAATGACA 2333
D 538 TTGTCGCGATTCCAGATTTCTATPACCTCATGCGCTTTGACAGGGCGTCTGCTGAACGACA 597
QY 2334 CCAAGACTTACAGCGCCGAGAGAAACCGTGTGTAATTTGGTTTTCAGCGGTACAGTGTACA 2393
D 598 CCAAACTTATCAGGCTGAGAGAGGTTCAAGGAGAGGTGGCTCTCTGCACTAGTGTGTACA 657
QY 2394 TGAGGGAATTCGGAGTGCACAGAGCAAGAGCTCTAAGTCTATGTTTATGTTATCATCG 2453
D 658 TGAAGGACCATCTCTGAGATCTCGGAAGAAGCTCTCAAAACATGTCTTACTACTGTATCG 717
QY 2454 ACAACGCACT 2463
D 718 ATATGCTCT 727

RESULT 3
LOCUS CF475347
DEFINITION RTW2_13_C12_g1_A021 Well-1 watered loblolly pine roots WM2 Pinus taeda cDNA clone RTW2_13_C12_A021 5', mRNA sequence.
ACCESSION CF475347
VERSION CF475347.1
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus 1 (bases 1 to 641)
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., Cannon, R., Owen, A. and Neale, D.
EST database from well-watered loblolly pine (Pinus taeda) roots
Other ESTs: RTW2_13_C12_b1_A021
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: JENREV (CAGGAACAGCTATGACC).
Location/Qualifiers
```

```
source 1..641
/organism="Pinus taeda"
/mol_type="mRNA"
/db_xref="CCONES"
/db_xref="taxon:3352"
/clone="RTW2_13_C12_A021"
/lab_host="DH108-T1 Phage-resistant E. coli"
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."
```

Query Match 9.9%; Score 267.4; DB 14; Length 641;  
Best Local Similarity 67.9%; Pred. No. 5.9e-37;  
Matches 389; Conservative 0; Mismatches 181; Indels 3; Gaps 1;

QY 1409 TTTCCGGGAGAAAGATAATGGAAGAGCTTAAGACCTTCTACTACAAATCATCTCCAAAT 1468  
D 3 TTTCCGGGAGAAAGCATATGAGGAGCAAACTCTGTACGGAAGGATCTCTGAAGCAT 62  
QY 1469 GCTCTTGCCTCAAAACCAACGCAATTTGATAAGTGGCTGTCAAGAAAGGATCTTCTCTGGAGAG 1528  
D 63 GCTCTGGAGGATGCGATGCTTTTGACAAATGGGCCCAATTAAGAAAGATCTTTCGGGCGAGAG 122  
QY 1529 GTGGAGTATGCTATAAGTATCGTGGCATAGAGTATGCCAAGATTCGAAGTGGAGGCAAGAGT 1588  
D 123 GTGGAGTATTCCTCAAAATATCTTGGCATAGAGTATGCCAAGGCTGGAGGCGCAGAGC 182  
QY 1589 TACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTA 1648  
D 183 TACGTTGAAATTTATGACCAAGAGATTTCTGGATCGGAAATTTTGTATTTGATGCA 242  
QY 1649 TATGTAGCAACGAAATAATTTGGAGCTGGCCAAATTTGACTTCAATATATGTCGAGGCC 1708  
D 243 AATGTTTTCAACGAAATAATTTGGAATTAGCAAAATTTGATTTCAATTCGACTGCAAGCT 302  
QY 1709 TTACACCAAGAGAGACTCAACATTTTGTGAGCTGGTGGGAGAGATCCGGAATTCATATGAT 1768  
D 303 TTACACAGACAAGAGCTTCAGGATATTCGGAGGTGGTGGAAATTCATCGGGTTTCTCAGAG 362  
QY 1769 CTTACATTTACCGCCAGCGCCCTGTG---GAAATGTATTTCTCAGTGGCGGTTAGTATG 1825  
D 363 CTGGCTTCACTCATCGGAAACGTGTGGCCGAAATATATTTCTCAGCAGCGTCTTTTG 422  
QY 1826 TTTGAGCCAGAAATTCGCTGTGTAGAAATTCCTATGCCAAGACTTCTTGGCTCGCAGTT 1885  
D 423 TTTGAACCAAGATTTGCTACTCTGCAGAGCGGTTTACACAAAACTGCTATTTTTCACCGTC 482  
QY 1886 ATTCTAGACGATCTTTACGACCCACCGATCTCTGGATGATCTTAATTTGTTCTCTGAA 1945  
D 483 CTATTAGATGATCTTTATGATCAAGCATGAAACTTTAGACAATCTTAACCTGTTTACAGAA 542  
QY 1946 GCGTTCGAAAGATGGGATATCTCTGTCTGGAT 1978  
D 543 TCTGTGAAAGATGGGATCTATCACTGGTAGAT 575

RESULT 4  
LOCUS AY109636  
DEFINITION Zea mays CL2355\_1 mRNA sequence.  
ACCESSION AY109636  
VERSION AY109636.1  
KEYWORDS GI:21213434  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoidae; Andropogoneae; Zea.		Db	853	TCCCAAGGGACATGATGCACAGGGTCCGACGTCCTCATCTCCACAG---NNNNNNNNNN	909
REFERENCE	1 (bases 1 to 2797)	Qy	906	AAGATGGTGGATCGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTAA	965
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.	Db	910	NN	969
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes	Qy	966	GCTCCCTGCATCTACTGCTTGTCTTTCATGCACACTCAAAACAAACGATCCCTCCACT	1025
JOURNAL	Unpublished (2002)	Db	970	NNNNNNNNNNCTACCGCTTACCGCTGATGCAACCGGTGNNNNNNNNNNNNNNNNNN	1029
AUTHORS	Coe,E.H.	Qy	1026	TTCTCACTTCGTGCTCAGCAAAATTTGGGACTAGTTCCTTCCATACCCACTTGATC	1085
REFERENCE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA	Db	1030	NNNNNNCAGGATGTCNNNNNTTCAACNNNNNAGTCCCAATGTTTATCCGTCGATC	1089
TITLE	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizeimap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.	Qy	1086	TATTTGAACGCTCTGGGCTGTCGATACAGATTGAAACGCTTGGGAATCGATCTATTCA	1145
JOURNAL	Direct Submission	Db	1090	TTTTGAGCACATCTGGGTGTGGATCGTTGGAGCGACTCGGATCTCCGCTACTTCC	1149
AUTHORS		Qy	1146	AGAAAGAAATCAAGAAATCTCTGGATTAGTTTATAGTTAGTGGACGCCGAAGAGCGC	1205
REFERENCE	Location/Qualifiers	Db	1150	AACGAGAGATTGAGCAGTGCATGGAATGTAACAGGACCTGG---ACTGAAGATGGGA	1206
TITLE	1. .2797	Qy	1206	TGGGATGGCAGATGCAATCCTATTCTGTGATGTCGATGACACTGCCATGGGTCTTAGAA	1265
JOURNAL	/organism="Zea mays"	Db	1207	TTTGCTGGGTAGGNN	1266
AUTHORS	/mol_type="mRNA"	Qy	1266	TCCTGAGACTTCATGGATACAATGTATCTTCAGATGTTCTCGAGAAATTTTCAGAGACA	1325
TITLE	/db_xref="MaizeDB:630936"	Db	1267	TACTAAGGCTACATGGATACATGTCTCTCAAGTGTGTTTAAAGAACTT---TGAGAAAG	1323
JOURNAL	/clone_lib="Maize Mapping Project/DuPont Consensus Library"	Qy	1326	AAGGAGACTCTTTTGTCTTTCGGCTCAACCAAAATGGTGTGACCGGATAATCTTAACC	1385
AUTHORS	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"	Db	1324	ATGGAGAGTTCTTTTGTGTTTGGGCCAATCGACTCAAGCCGTCACCTGGGATGTATAACC	1383
REFERENCE		Qy	1386	TTTATAGATGTTCAAAAGTATGTTTTCGGGAGAAAAGATAATGGAAAGCTTAAGACCT	1445
TITLE	Query Match	Db	1384	TCACAGAGCTCTCAGATAAGTTTTCAGAGGAGAGGATGATTTGTCATCGTGTAGGGTTT	1443
JOURNAL	Best Local Similarity	Qy	1446	TCACTACAAATCATCTCCAAATGCTCTTGGCCAAAACAAACGCAATTTGATAGTGGGCTG	1505
AUTHORS	Matches	Db	1444	TCCTGATGAGTTTCNN	1503
TITLE	Conservative	Qy	1506	TCAGAGAGATCTTCTCGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAGTA	1565
JOURNAL	0; Mismatches	Db	1504	TTGCCAAGGATCTACCTCGGAGGTGCAATATACACTAGACTTCCCTTGGTATGCAAGCT	1563
AUTHORS	686; Indels	Qy	1566	TGCCAAGATTGGAGGCAAGAGTTACATAGAGCAAT---TTGGATCAAAATGATGTCTGGC	1622
TITLE	21; Gaps	Db	1564	TGCTCGTGTAGAGGCAAGAACCTATCTAGATCAATATGGTGGTAAAGATGACGTTTGGGA	1623
REFERENCE		Qy	1623	TGGGGAAGACTGTATATAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGCCA	1682
TITLE	9.8%; Score 264.2; DB 11; Length 2797;	Db	1624	TTGGAAGACACTCTACAGATGCTCTTGTGTAATACGNNNNNNNNNNNNNNNNNNNA	1683
AUTHORS	45.4%; Pred. No. 1.3e-36;	Qy	1683	AATTGACTTCAATATGTTGTCAGGCTTACACAAAAGGAGACTCAACACATGTCTAGCT	1742
JOURNAL	60; Mismatches 803; Indels	Db	1684	TAGGGATTCAACCATTCGCAAGCTCTGCATCAGCTTGAGTGTAAATGGGCTGCAACGCT	1743
TITLE		Qy	1743	GGTGGAGAAATCGGGATTCAATGATCTTACATTCACCCGCCAGCGGCTGTGGAAATGT	1802
REFERENCE		Db	1744	GGTACAAGGATAATTGCTTTGACGCTTTTGGAGTAGAAACCAAGATGTTTAAAGATCTT	1803
TITLE		Qy	1803	ATTTCTCAGTGGCGTTAGTATGTTTGGAGCAGAAATTCGCTGCTTGTAGAAATTCGCTATG	1862
AUTHORS		Db	1804	ACNNNNAGTGTGCTTGCNNNNNNGAACCTAGCCGCTGCTGCTGAGCGGCTTGCGATGGG	1863
JOURNAL		Qy	1863	CCAAGACTTC	1872
REFERENCE		Db	1864	CTAGAACGTC	1873
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE		Db			
TITLE		Qy			
AUTHORS		Db			
JOURNAL		Qy			
REFERENCE					

LOCUS BF517045 476 bp mRNA linear EST 07-MAY-2003  
 DEFINITION NXSI\_008\_E08\_F NXSI (Nsf Xylem Side wood Inclined) Pinus taeda cDNA clone NXSI\_008\_E08\_5' similar to Arabidopsis thaliana sequence At4g02780 gsf-kaurene synthetase A - like protein see <http://mips.gsf.de/proj/thal/db/index.html>, mRNA sequence.  
 ACCESSION BF517045  
 VERSION  
 KEYWORDS  
 SOURCE Pinus taeda (loblolly pine)  
 ORGANISM  
 Pinus taeda  
 Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 476)  
 Sederoff, R.  
 Molecular Basis of Wood Formation in the Pine Megagenome  
 Unpublished (2000)  
 CONTACT: Sederoff, Ron  
 Forest Biotechnology  
 North Carolina State University  
 840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh, NC 27695, USA  
 Tel: 919 515 7800  
 Fax: 919 515 7801  
 Email: [ron.sederoff@ncsu.edu](mailto:ron.sederoff@ncsu.edu), [jerry.johnson@ncsu.edu](mailto:jerry.johnson@ncsu.edu)  
 Please see <http://web.ahc.umn.edu/biodata/nsfpine/> for further information.  
 Seq primer: T3.

# FEATURES

Location/Qualifiers  
 1..476  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="Coastal plain loblolly pine from North Carolina"  
 /db\_xref="taxon:3352"  
 /clone="NXSI\_008\_E08"  
 /tissue\_type="Xylem"  
 /cell\_type="Side"  
 /dev\_stage="Juvenile"  
 /lab\_host="XLI-Blue"  
 /clone\_lib="NXSI (Nsf Xylem Side wood Inclined)"  
 /note="Vector: Bluescript SK; Site 1: Eco RI; Site 2: XhoI; The library is from early (spring) wood, taken from three six-year old trees (three different genotypes), in the juvenile phase. These trees were induced to form side wood by bending to a 45 degree angle and tying them to the ground. Differentiating xylem was harvested from the sides of the inclined stems, and a mixture of all three genotypes was used for the library. oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI Bluescript SK vector arms. NOTE: The sequences contain a 'cDNA adapter' between the EcoRI site and the start of the EST. The adapter sequence is 'AATTGGGACGAG'."

## ORIGIN

Query Match 8.6%; Score 231.4; DB 10; Length 476;  
 Best Local Similarity 66.4%; Pred. No. 1.3e-30;  
 Matches 316; Conservative 0; Mismatches 160; Indels 0; Gaps 0;  
 643 CATTCAAGTGCAGAAAGGGTTGAGTTGTGTGAGAAACACATGGAGAAATGAAGACGA 702  
 Db 1 CACACAAGTACAAAAAGGGATTGAATTTTCAAGACACAAAGCTGGAAAAATAGAGAGGA 60  
 703 AGCTGACATACACAGCCCAAGTGGATTGAGGTCTGTTCCTGCAATGTTAGATGAAGC 762  
 Db 61 GGCTGATAGCCACAGCCCAAGTGGATTGAAATAGTATTTCTGCAATGCTAAAGAGAC 120  
 763 AAAAAGCTTGGGATTTCTCTTATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 822  
 Db 121 TAAAGCTTAGTCTGGCTCTGCTCTGCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
 823 GCGCAGAAAAAGCTTCAAAAGATTCCTCTCAATGTTCTTCTATACCATCATCAGACGGCTT 882  
 Db 181 GCGGGAGGCTAAGCTTCAAGAGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240

883 GCTCTACTCTCTGGAGGTTTTCAGATGTGTGGACTGGCAAGAGATCACAAATCTTCA 942  
 Db 241 ACTGTATTCTTTTGAAGNCCTGCAAGAAATAGTAGATGGGAGAAAATAATGAACATCTCA 300  
 943 ATCAAGAGACGGATCAATTTTAAAGCTCCCTGTCATCTACTGCTGTGTCTTCATCATCACAC 1002  
 Db 301 ATCCANNNTGGTCAATTCCTCAGCTCGCCAGCATCTACAGGGCTGTGTTCATCGGTAC 360  
 1003 TCAAAACAAACGATGCTCCACTTCTTCAACTTCGTGCTCAGCAAAATTTGGCGATACGT 1062  
 Db 361 AGGGAACAAAATAATGCTGNNNNNTCTTGAAATTTGTCTTGAAGAAATTCGGAACCATGT 420  
 1063 TCCTTGGCATTTACCCACTTGATCTATTGAAAGCCTCTGGGCTGTCCGATACAGTTG 1118  
 Db 421 TCCTTGTCACTATCGCTTGANNCTTTTGAACGNNNNNGCCGTTGATCTGTNG 476

## RESULT 6

LOCUS CF397293 725 bp mRNA linear EST 29-AUG-2003  
 DEFINITION RTDS3\_2\_A09\_g1\_A022 Drought-stressed loblolly pine roots DS3 Pinus taeda cDNA clone RTDS3\_2\_A09\_A022 5', mRNA sequence.  
 ACCESSION CF397293  
 VERSION  
 KEYWORDS  
 SOURCE Pinus taeda (loblolly pine)  
 ORGANISM

## REFERENCE

1 (bases 1 to 725)  
 Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J., and Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., and Neale, D.  
 An EST database from drought-stressed loblolly pine (Pinus taeda) roots  
 Unpublished (2003)  
 Other ESTs: RTDS3\_2\_A09\_b2\_A022  
 CONTACT: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: [mmpratt@uga.edu](mailto:mmpratt@uga.edu)

## TITLE

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGAAACAGCTATGACC).

Location/Qualifiers  
 1..725  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONES"  
 /db\_xref="taxon:3352"  
 /clone="RTDS3\_2\_A09\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Drought-stressed loblolly pine roots DS3"  
 /note="Vector: pSli180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from drought-stressed loblolly pine (Pinus taeda) roots. Water was withheld from ramet clones until predawn needle water potential reached -1.75 MPa. On day 7 roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSli180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

## FEATURES

source  
 Query Match 8.5%; Score 228.8; DB 14; Length 725;  
 Best Local Similarity 58.3%; Pred. No. 3.2e-30;  
 ORIGIN  
 Query Match  
 Best Local Similarity

Matches 420; Conservative 0; Mismatches 297; Indels 3; Gaps 1;		Seq primer: T3.		Location/Qualifiers	
1533 AGTATGCTATAAAGTATCCGTGGCATAGAAATGCCAAGATTCGAGGCAAGAAGTTACA 1592		1. 583		/organism="Pinus pinaster"	
Db				/mol_type="mRNA"	
2 AATACGCCCTGAATCTCTTGGCATGCGAGTGTCCCAAGAGTGGAGCAAAAAGATACT 61				/db_xref="taxon:71647"	
1593 TAGAGCAATTTGGATCAATATGATGCTGGCTGGGAAGACTGTGTATAGATGCTATATG 1652				/clones="RS62E09"	
Db				/tissue_type="root"	
62 GTCAAGTGTATATCCGGATATGCGATGTTTAGCAAAAAGTGTTCACAGCTGTCCACG 121				/dev_stage="6 weeks old seedling"	
1653 TGAGCAACGAAATATTTGGAGCTGGCCAAATTCGACTTCAATATGGTGCAGGCGCTTAC 1712				/lab_host="SOLR"	
Db				/clone_lib="RS"	
122 TGAACACGAAAGCTCCTGGATTTGGCAAAATTAGATTTCAACATGTGCCAGTCCATCC 181				/note="vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A three weeks drought stress treatment was applied by lowering the osmotic potential of the nutrient solution to -0.45 MPa using 170 g/l of polyethylene glycol as an osmoticum. A mixture of genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-exclused to form a phagescript phagemid"	
1713 ACCAAAGGAGACTCAACACATGTGCTAGCTGTGGAGAGAAATCGGATTCATATGATCTA 1772					
Db					
182 ACCAAGAGAGATGAAGAGTGTACAGCTGGTTTAAAGATTCGGGGTTGCCCTCTATTCA 241					
1773 CATTCACCGCCAGCGGCTGTGGAATGATTTCTCAGTGGCGTTAGTATGTTTGAGC 1832					
Db					
242 CTTTGCACGGGAGAGACCGCTGGAAATCTCTCTTGATAGCCGCTGTACGTATGAGC 301					
1833 CAGAAATTCGCTGCTGTAGAAATGCTATGCCAAGACTTCTTGCTCGCAGATTAATCTAG 1892					
Db					
302 CCCAGTATGCCAAATGCAGATTTCTCTTACAAAAGTGGCGTCTTGCAGACTGTTCTGG 361					
1893 ACGATCTTACGACACCCAGGATCTCTGGATGATCTTAAATTTGTTCTCTGAAGCGGTCC 1952					
Db					
362 ACGATATGATGACATTTATGGACCCCTCGATGAATTTGAAGCTGTTTCACTGAGGCTGTA 421					
1953 GAAGATGGGATATCTCTGTGCTGGATAGCGTTCCGGGATAATCATGTTTGAAGTTTGTCTCC 2012					
Db					
422 GAAGATGGGACCTCTCTTACGGAACCCCTCCAGACTAT---ATGAAGCTATGTTTACA 478					
2013 TAGGGCTGTACACACAGTGAATGGATTTGGAAGAAGTGAAGTCAAGGACAGGCGGTG 2072					
Db					
479 AAATCTATTATGACATAGTTTACAGATGTCCTGGAGGCGCTGGAGGAGAGAGGACAGGACGTG 538					
2073 ATGTGCTGGGCTATCTTCGAAAGTATGGAGGCGTTCGTCGATCGTATACCAAGAAG 2132					
Db					
539 AATTAGTTAGCTTTTCAGAAAGGATGGAGGATATCTCTCTGGGTTATTATGAAGAG 598					
2133 CCGAATGGTCGGCAGCAAAATGATGTGCCGACATTCACGAATATGTGGAAGTGGCAAG 2192					
Db					
599 CTGAATGGCTAGCAGCTGAGTATGTGCCAGTGTGGACGAATACATAAAGAAATGGAATCA 658					
2193 TGTCATAGCACTTGCACAGCTGTACTAAACTCAATCTTTTTCACCTGGAGAAATTACTTC 2252					
Db					
659 CTTCCATTTGGCCAACTATCTCTCTTTGAGTGAGTATTTGATATATGAGGAGCACTCC 718					
RESULT 7		583 bp		mRNA	
BX681656		RS		Pinus pinaster cDNA clone RS62E09, mRNA sequence.	
LOCUS					
DEFINITION					
BX681656					
VERSION					
BX681656.1					
KEYWORDS					
EST.					
SOURCE					
pinus pinaster					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.					
1 (bases 1 to 583)					
Frigerio,J. and Plomion,C.					
IDENTIFICATION					
of water-deficit responsive genes in Maritime pine					
(Pinus pinaster Ait.) using an EST approach					
UNPUBLISHED (2002)					
CONTACT: Frigerio JM					
Genetique et Amelioration 69					
INRA					
route d'Arcachon 33612 Cestas CEDEX France					
Email: Frigerio@pierrot.inra.fr					
Email: Frigerio@pierrot.inra.fr					
JOURNAL					
COMMENT					
RESULT 8		634 bp		mRNA	
CF397946		RTDS3 23 G06.g1 A022 Drought-stressed loblolly pine roots DS3 Pinus		linear	
LOCUS					
DEFINITION					
taeda cDNA clone RTDS3 23 G06_A022 5', mRNA sequence.					
CF397946					
ACCESSION					

Query Match 7.9%; Score 214.8; DB 13; Length 583;  
Best Local Similarity 63.6%; Pred. No. 9.7e-28;  
Matches 361; Conservative 0; Mismatches 202; Indels 5; Gaps 2;

ORIGIN

1695 ATATCGTCAGGCGCTTACACAAAGGAGACTCAACACATTTCTCAGCTGGTCGAGAGAAAT 1754  
Db 1 ATCGTGTGCAATTTTATACACAGACAGAGAGCTTCAAGATATCGGAGG--TGGTGGATCAT 58  
1755 CCGGATTCATCATCTTACATTCACCCGCGCGCTGTGGAAATGTATTTCTCAGTGG 1814  
QY 59 CCGGTTTCTCAGAGCTGGGCTTCTCGGGAAGGTGCTGAAATATATTTTTCACAG 118  
Db 1815 CCGTTAGTATGTTGAGCCAGAAATTCGCTGCTTTAGAAATTCGCTATGCCAGACTTCTT 1874  
QY 119 CGTCAATTTTGTGTGAGCCAGAGTTCTGCTACCTGCGGAGCCGTTTACACAAAACCTCTA 178  
Db 1875 GCCTCGAGTTATTTAGACGATCTTACGACACCCAGCGATCTCTGGATGATCTTAAAT 1934  
QY 179 ATTTACGGCTCATATTAGATGATCTTTATGAGCTCATGGAATTTGGACATCTCAAGT 238  
Db 1935 TGTCTCTCAAGCGGTCGGAAGATGGATATCTGTGTGTGATAGCGTTGGGATATC 1994  
QY 239 TGTTCAGAACTGTGAAAGATGGATCTATCTCTGTGTAGTCAATATGCCCAA--G 295  
Db 1995 AGTTGAAGTTTCTTCTAGGCTGTACACACAGTGAATGATTTGGAAGATGGAC 2054  
QY 296 ACATGAAATATGTTTAAAGGGTTTATCAATATCTTTTAAATGAAATCGCTGAAGAGGGC 355  
Db 2055 TCAAGGAACAAGCGCTGATGTCTGGGCTATCTTCGAAAGATGAGGAGGCTTGTCTG 2114  
QY 356 GGAAGAGGCGGCGCTGATGTCTAGTTTACATTCAAAAGTTTGGAGGTCAGCTAG 415  
Db 2115 CATCGTATACAAAGAGCGGAATGTGCGCAGCAAAAGTATGTGCGACATTCACCAAT 2174  
QY 416 GAGCGTATACAAAGAGCAGAAATGGTGTGACGTAGATACGTGCGCGCTCTCTATGATGAAT 475  
Db 2175 ATGTGGAATGCCAAAGTGTCCATAGCACTTGCAGCAGTCGTACTAAATCAATCTTTT 2234  
QY 476 ATATAGGAACCGAGTGTTCATAGCATTTGGACATTTGGACAGTGGTTCATCAGCGCTCTTT 535  
Db 2235 TCACCTGGAGAAATTAATCTTCTCGATTACAT 2262  
QY 536 TCACCGGGAGATCTTACACAGTATAT 563

RESULT 8  
CF397946 634 bp mRNA linear EST 29-AUG-2003  
LOCUS  
DEFINITION RTDS3 23 G06.g1 A022 Drought-stressed loblolly pine roots DS3 Pinus  
taeda cDNA clone RTDS3 23 G06\_A022 5', mRNA sequence.  
CF397946  
ACCESSION





Query Match 7.6%; Score 204.4; DB 14; Length 725;  
 Best Local Similarity 58.9%; Pred. No. 5.9e-26;  
 Matches 371; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

1623 TGGGGAAGACTGCTATAGATGCTATATGTGAGCAACGMAAAATATTGGAGCTGGCCA 1682  
 Db 5 TAGCAAAAGTGTTCACAGCTGCCCTAGTGNACACGMAAAGCTTCAGAGTTGGCA 64

1683 AATTGGACTTCAATATGTTGCGAGGCTTACACCAAAAGAGACTCAACACATGTCAGCT 1742  
 Db 65 AATTAGATTCAACATTTGCCAGTCCATCCACCAAGAGAGATGAAAAATGTTACCAGCT 124

1743 GGTGAGAGAAATCGGATTCATATGATTCATATCCCGCCAGCGCTGGGAAATGT 1802  
 Db 125 GGTTTAAAGATTCGGGTTGCCCTTATTCACCTTTGACGGAGAGCGCTGGAAATCT 184

1803 ATTCTCAGTGGGTTAGTATGTTTGGAGCCAGAAATTCGCTGCTTGTAGAAATGCTATG 1862  
 Db 185 ACTTCTGATAGCGCTGTAGTATGAGCCCGCAGTATGCCAAATGCAGATTCCTCTTTA 244

1863 CAAGACTTCTTCCCTCGCAGGATTTCTAGACGATCTTTACACACCCAGGATCTCTGG 1922  
 Db 245 CAAAAGTGGCGTGTTCGCTGACTGTTCTGAGCAGATATGATGACACTTATGGGACCTTG 304

1923 ATGATCTTAAATGTTCTCTGAAGCGTCCGAGATGGGATATCTCTGTGCTGGATAGCG 1982  
 Db 305 ATGAATGGAAGCTGTTCACTGAGCGCTGAGAGATGGGACCTCTCTTATACAGAAACC 364

1983 TTCGGGATTAATCAGTTGAAAGTTTGTCTCTAGGCTGTACACACAGTGAATGATTTG 2042  
 Db 365 TTCAGACTAT--ATGAACTATGTTACAAATCTATTATGACATAGTTCACGAAGTGG 421

2043 GAAAGATGACTCAAGGACAAAGCGGTGATGCTGGCTATCTTCGAAAGATGAGG 2102  
 Db 422 CTGGGAGGAGAGAGGACGAGCGGTGAATTTGGTTAGCTTTTTCAGAAAGGATGG 481

2103 AGGCTTGTCTCGCATCGTATACCAAGAGCGGATGTTGGCGAGCAAAAGTATGTCGCA 2162  
 Db 482 AGAATTAATCTGAGTATATGAGAGAGCTGAATGGTTAGCACTGAGTATGTCGCA 541

2163 CATTCACGAATATGTTGAAATGCAAAAGTGTCCATAGACATTTGGCAGAGTGTACTAA 2222  
 Db 542 GCTTGAACAGTATACATCAAGACGGAATCACTTCCATTTGGCCCAACGTATCTCTGTTGA 601

2223 ACTCAATCTTTTCACTGGAGATTAATCTTC 2252  
 Db 602 GTGGAGTATGATTAATGGAGGACCACTCC 631

RESULT 10  
 CF474786  
 LOCUS  
 DEFINITION  
 CF474786 700 bp mRNA linear EST 05-SEP-2003  
 taeda cDNA clone RTWw2\_7\_Bil\_A021 5', mRNA sequence.

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

Pinus taeda (loblolly pine)  
 Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 700)  
 Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W., Dean, J.,  
 Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
 Neale, D.  
 An EST database from well-watered loblolly pine (Pinus taeda) roots  
 Unpublished (2003)  
 Other ESTs: RTWw2\_7\_Bil.bi.A021  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of  
 Forestry, University of Georgia; plant material prepared at the  
 University of Florida; sequencing done in the Laboratory for  
 Genomics and Bioinformatics, University of Georgia. Sequence ends  
 have been trimmed to exclude vector and regions below Phred quality  
 16. Three-prime sequences are presented as their reverse complement  
 and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGGAACAGCTATGACC).

FEATURES

source

1..700  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONES"  
 /db\_xref="taxon:3352"  
 /clone="RTWw2\_7\_Bil\_A021"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots Ww2"  
 /note="Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The  
 library was prepared from polyA+ RNA from loblolly pine  
 (Pinus taeda) roots watered to pot capacity every other  
 day. Pre-dawn water potential remained -0.3 MPa +/-0.1.  
 Roots were harvested for RNA isolation. Double-stranded  
 cDNA was cloned unidirectionally into pSL180. Inserts  
 excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 7.1%; Score 190.8; DB 14; Length 700;  
 Best Local Similarity 55.6%; Pred. No. 1.4e-23;  
 Matches 388; Conservative 0; Mismatches 307; Indels 3; Gaps 1;

1787 CGGCTGTGGAATGTATTTCTCAGTGGCGGTAGTATGTTTGGAGCAGAAATTCGTCT 1846  
 Db 2 CGTACGTGGATACTACACTTTAGCAGCTTGCAATGATCTTAACATTTCTGCG 61

1847 TGTAGAAATGCTATGCGCAAGACTTCTGCTCGCAGTATTTCTAGACGATCTTTACGAC 1906  
 Db 62 TTTTCGACTAGGATTTGGTAAATAAGTCATATGATCAGATTTCTGACGATATCTAGAC 121

1907 ACCCAGGATCTTGATGATCTTAAATGTTCTCTGAAGCGGTCGGAAGATGGGATATC 1966  
 Db 122 ACCTTCGGAACATGAGGAGCTCGACTTTAAACGCGAGCTTTAAGAGATGGGATCCG 181

1967 TCTGTCTCGATAGCGTTCTGGGATATCAATGATGAAAGTTTGTCTTAGCGCTGTACAC 2026  
 Db 182 TCTTCGATAGAGTGTCTCCAGATTAT---ATGAAGAGAGTGTACATGGCGGTTTACGAC 238

2027 ACAGTAATGGATTTGGAAGATGGACTCAAGGAACAAGCGCTGATGTGCTGGGCTAT 2086  
 Db 239 ACATCAAGAAATGCGAGAGGCGGAGAAATTCAGGCTGGATACAGTCAGCTAT 298

2087 CTTTCAAAAAGTATGGAGGCTTGTCTCGCATCGTATACCAAGAGCCGAATGGTGGCA 2146  
 Db 299 GCTCGAAATCTTGGGAGGCTTTATTTGGTGTATATACAGAAAGCAAGTGGATTTC 358

2147 GCAAGTATGTCCGACATTCACGAATATGTGGAATGCCAAAGTGTCCATAGCACTT 2206  
 Db 359 AGTGGTTATCTTCCCGTTCGAGATACCTCGAATGGGAGGTGACGCTTCGGCTCT 418

2207 GCACAGTCTACTAAACTCAATCTTTTTCACCTGGAGAAATTAATCTCTGATTAATTTTA 2266  
 Db 419 CGCATACACAGCTCGAACCCATGCTGACTTTGGGGTTTCTCTTCGCGCTCGAATCTG 478

2267 CAGCAAGTATAGACTTCGGTCCAAATTTCTGATCTTGTGTCTTTGATGACGACTAATC 2326  
 Db 479 CAGGAATTTGACTTTCATCGAAATTCATGATATGTCATCTTCTTCGACTGAAA 538

2327 AATGACACCAAGACTTACCGAGCGGAGAGAAACCGTGTGTAATTTGGTTTCCAGGCTACAG 2386  
 Db 539 GGTGACACTCAATGCTACAAGGCTGACAGGCGGCTGGAGAGAGAGTTCGGCGGTATCG 598

QY 2387 TGCTACATGAGGGAATCCGGAGTGCACAGAGGAAGCTCTAAGTCAATGTTTATGCT 2446  
 Db 599 TGTATATGAAGACCATCTCGAATACAGAGGAAGATGCTGTCAATCAAGTCAATGCT 658  
 QY 2447 ATCATCGACACGACCTGAAGGAATGAATGGGAGTT 2484  
 Db 659 ATGTCGATTAACCAAGGAACCTGAATGGGAGTT 696

RESULT 11  
 LOCUS CF477103  
 DEFINITION RTW3\_5\_A06.g1\_A022 Well-watered loblolly pine roots WW3 Pinus taeda cDNA clone RTW3\_5\_A06\_A022 5', mRNA sequence.  
 ACCESSION CF477103  
 VERSION 1  
 KEYWORDS EST.  
 SOURCE CF477103.1 GI:34505972  
 ORGANISM Pinus taeda (loblolly pine)

Pinus taeda (loblolly pine)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 745)  
 Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Dean,J., Gebremedhin,M., Dervinis,C., Martin,T., White,T., Davis,J. and Neale,D.  
 An EST database from well-watered loblolly pine (Pinus taeda) roots

Unpublished (2003)  
 Other ESTs: RTW3\_5\_A06.bi\_A022  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGGAACAGCTATGACC).  
 Location/Qualifiers

## FEATURES

source  
 1. 745  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONES"  
 /db\_xref="taxon:3352"  
 /clone="RTW3\_5\_A06\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots WW3"  
 /note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 Mpa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

## ORIGIN

Query Match 7.0%; Score 190.4; DB 14; Length 745;  
 Best Local Similarity 54.8%; Pred. No. 1.7e-23;  
 Matches 399; Conservative 0; Mismatches 326; Indels 3; Gaps 1;  
 QY 1852 AATTGGCTATGCCAAGACTTCTGCTCGCAGTATTCTAGACCATCTTTACGACACCA 1911  
 Db 1 ACTGGCTTTGCCAAATCTGTCATATACATCACCCTTGTGATATGTCAGACACTT 60  
 QY 1912 CGGATCTCTGATGATCTTAATGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGT 1971  
 Db 61 CGGAACACTCGACGCTCGACTTTCACAGCTGCAATTAAGAGATGGGATCGTGGC 120

QY 1972 GCTGGTAGCTTCGGGATAATCAGTTGAAAGTTTGCTTCTCTAGGGCTGTACACACAGT 2031  
 Db 121 GACAGAGTGCCTTCCAGATATATGAAAGGAGTTTACATGATAG---TTTACAACTGT 177  
 QY 2032 GAATGGAATTTGGAAAAGATGGAATCTCAAGGAACAAGCCGCGTGAATGCTGGGCTATCTTCG 2091  
 Db 178 AATGGAATGCTCTCAGGAGCGACAGAGGCTCAAGGCCGAGACACGCTCAACTATTGTCTG 237  
 QY 2092 AAAAGTATGGGAGGGCTTGTCTCGATCGTATACCAAGAAGCCGAATGTCGGCAGCAAA 2151  
 Db 238 ACAGGCTTGGGAGGAATATATTGATTCTGATATATCAAGAAGCAAAAGTGGATCGCCAGTGG 297  
 QY 2152 GTATGTGCCACATTCACGAATATGTGGAAATATGCCAAAGTGTCCATAGACCTTGGAC 2211  
 Db 298 TGAGGTGCCAACATTTGAGGAGTACTACGAGAAGCGGAAATTAGCTCTGTCTCATCGGT 357  
 QY 2212 AGTCGTACTAAACTCAATCTTTTCTACTGGAGAATTAATCTTCTCGATTACATTTTACAGCA 2271  
 Db 358 ATCGGCATTGCAACCCCAATCTGACGACCGACATCCCTTTCTTGAGCAGCTCTCAAGGA 417  
 QY 2272 AGTAGACCTTGGTCCAAATTTCTGCAATCTTGTGTCTTTGACTGGAGCAGCTAATCAATGA 2331  
 Db 418 AGCTGACATTTCCATCGAAGCTCAATGACTTTGGCATCTGCCATTTCTTCCGATTACGAGGA 477  
 QY 2332 CACCAAGACTTACCAGSCCGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTACAGTGTCTA 2391  
 Db 478 TAGCGCTGCTACAGCGGACAGGGCCCTGTGGAAGAAGCTTCGTGTATATCTTGTTA 537  
 QY 2392 CATGAGGGAATCCGAGTGCACAGAGGAAGAGCTCTAAGTCATCTTTTATGTATCAT 2451  
 Db 538 TATGAAGACAACTCTGGAGCAACGGAAGAAGATGCTCTCAATCATATCAACGCCATGAT 597  
 QY 2452 CGACACGCACTCGAAGGAATGGAATGGGAGTGGCCCAACCCAGCGAGCAATGCCCATTT 2511  
 Db 598 CAGTGTATTAATTAAGGATTAATTTGGGAGCTTCTCAACCAACAGCAGCGGTCCCAT 657  
 QY 2512 GTCTGTGAGAAGACTGCTGTTCAACTGCAAGAGTATGATCGAGCTGTTTATATGTACAG 2571  
 Db 658 ATCTGCCAAAACATGCTTTTGACATTAGCAGGCTTTCCATTATGGGTACAAATATCG 717  
 QY 2572 AGATGGCT 2579  
 Db 718 AGATGGCT 725

## RESULT 12

CF401916  
 LOCUS RTW1\_15\_B05.g1\_A015 Well-watered loblolly pine roots WW1 Pinus taeda cDNA clone RTW1\_15\_B05\_A015 5', mRNA sequence.  
 DEFINITION CF401916  
 ACCESSION CF401916.1 GI:34360333  
 VERSION 1  
 KEYWORDS EST.  
 SOURCE Pinus taeda (loblolly pine)  
 ORGANISM Pinus taeda (loblolly pine)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.  
 1 (bases 1 to 696)  
 Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Dean,J., Gebremedhin,M., Dervinis,C., Martin,T., White,T., Davis,J. and Neale,D.  
 An EST database from well-watered loblolly pine (Pinus taeda) roots  
 Unpublished (2003)  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for

## TITLE

## JOURNAL

## COMMENT

Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

## FEATURES

source

1..696  
Location/Qualifiers  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTW1 15 B05 A015"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/clone\_lib="Well-watered loblolly pine roots W1"  
/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

## ORIGIN

Query Match 7.0%; Score 188.8; DB 14; Length 696;  
Best Local Similarity 55.7%; Pred. No. 3.2e-23;  
Matches 363; Conservative 0; Mismatches 302; Indels 3; Gaps 1;  
QY 1828 TGAGCCAGAAATTCGCTGCTGTAGATTGCTTATGCCAGAGCTTCTTGCTCGCAGTTAT 1887  
DB 12 TGATCTTAACAACTTTGGCTTTCGACTAGATTGGTAAATAAGTCATATATACGAT 71  
QY 1888 TCTAGACGATCTTTACGACACCCAGGATCTCTGGATGATCTTAAATTTCTTGAAGC 1947  
DB 72 TCTCGAGATATCTACGACACCTTCGGAACAATGGAGGAGCTCGAACTCTTAACCGCAGC 131  
QY 1948 GGTCCGAAGATGGGATATCTCTGCTGTGATAGCGTTCCGGTAAATCACTTGAAGTTTG 2007  
DB 132 GTTTAAGAGATGGGATCCGCTCTTCGATAGATGTCTTCCAGATTAT--ATGAAGAGAT 188  
QY 2008 CTTCTTAGGGCTGTACACACAGTGAATGGATTGTGAAAAGATGGACTCAAGGAACAAGG 2067  
DB 189 GTACATGGCGGTTTACGACAACTCAACGAAATGGCAGAGGCGCAGAAATTCAGG 248  
QY 2068 CCGTGATGTCTGGGCTATCTTCGAAAAGPATGGAGGGCTTGCTCGCATCGTATACCAA 2127  
DB 249 CTGGGATACAGTCAGCTATGCTCGAAAATCTTGGAGGGCTTTTATTTGGTCTTATATACA 308  
QY 2128 AGAAGCGAATGTTCGGCAGCAAGATGTGCGGACATTCACGCAATATGTGGAATGC 2187  
DB 309 AGAAGCAAGTGGATTTCAGATGGTTATCTTCCACGTTTCGACGATGACCTCGAGATGG 368  
QY 2188 CAAAGTGTCCATAGCACTTCGACAGTCGTACTAAACTCAATCTTTTCACTGGAGAAAT 2247  
DB 369 GAAGTCAAGTTCGGCTCTCGCATACCAACGCTCGAACCCATCTGACTTTGGGGTTTC 428  
QY 2248 ACTTCTGATTAATTTTACGCAAGTAGACCTTCGGTCCAAATTTCTGCAATCTTGTC 2307  
DB 429 TCTTCGGCTCGAATCTCGAGGAAATTTGACTTTCCATCGAAATTCAAATGATTGTATG 488  
QY 2308 TTTGACTGGAGCACTAATCAATGACACCAAGACTTACCGGCGGAGAAACCGTGTGA 2367  
DB 489 TGGCATCTTCGACTGAAGGTGACATCAATGCTACAGCTGACAGGCGCGTGAGA 548  
QY 2368 ATTGGTTTCCAGCGTACAGTGTCTACATGAGGGAATTCGGAGTGCACAGAGGAAGGC 2427  
DB 549 AGAAGCTTCGGCGGATCGTGTGTATATGAAAGACCATCTCTGGAATTAACAGAGGAATGC 608  
QY 2428 TCTAAGTCATGTTATGGTATCATCGACAGCACTGAAGGAATTTGAATTTGGAGTTGGC 2487  
DB 609 TGTCAATCAAGTCAATGCTATGGTCGATACTTAACCAAGGAACCTGAATTTGGAGTTACT 668  
QY 2488 CAACCCAGCAGCAATGCCCATTTGTGT 2515

DB 669 TAGACCCGACAGCGGTGTTCCTCCATCTCT 696

## RESULT 13

LOCUS BX677599  
DEFINITION BX677599 RN Pinus pinaster cDNA clone RN41G08, mRNA linear EST 28-OCT-2003  
ACCESSION BX677599  
VERSION BX677599.1 GI:38011551  
KEYWORDS EST.  
SOURCE Pinus pinaster  
ORGANISM Pinus pinaster

REFERENCE  
AUTHORS Frigerio J. and Plomion C.  
TITLE Identification of water-deficit responsive genes in Maritime pine (Pinus pinaster Ait.) using an EST approach  
JOURNAL Unpublished (2002)  
COMMENT Contact: Frigerio JM  
Genetique et Amelioration 69  
INRA

route d'Arcachon 33612 Cestas CEDEX France  
Email: Frigerio@pierrot.inra.fr  
Email: Frigerio@pierrot.inra.fr  
Seq primer: T3.  
Location/Qualifiers  
1..581  
/organism="Pinus pinaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:71647"  
/clone="RN41G08"  
/tissue\_type="root"  
/dev\_stage="6 weeks old seedling"  
/lab\_host="SOLR"  
/clone\_lib="SOLR"

## FEATURES

source

1..581  
/organism="Pinus pinaster"  
/mol\_type="mRNA"  
/db\_xref="taxon:71647"  
/clone="RN41G08"  
/tissue\_type="root"  
/dev\_stage="6 weeks old seedling"  
/lab\_host="SOLR"  
/clone\_lib="SOLR"  
/note="Vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A mixture of genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBluescript phagemid"

## ORIGIN

Query Match 6.9%; Score 186; DB 13; Length 581;  
Best Local Similarity 64.9%; Pred. No. 1.1e-22;  
Matches 292; Conservative 0; Mismatches 155; Indels 3; Gaps 1;  
QY 1734 TTGTCAAGTGTGGAGAGATCGGATTCATATGATCTTACATTCACCCCGCAGCGGCTG 1793  
DB 135 TTCAAAGTGTGGAGATCATCCGTTTCGCGAGCTCAATTTACACGCGACCGTGTGA 194  
QY 1794 TGAATGATTTTCTCAGTGGCGGTAGTATGTTGAGCCAGAAATTCGCTGCTGTAGAA 1853  
DB 195 CTGAATATTTTCTCAATCGGCTCTCCCTGTTTGAACAGAGCTCGCTTCGTGTAGAG 254  
QY 1854 TTGCTTATCGCAAGACTTCTTGCTCGCAGTTATCTTAGACGATCTTTAGCACACCCACG 1913  
DB 255 CCAATTATAGAAATCTACCACTTGTGATTTGTATTTAGACGACCTATATGACCATG 314  
QY 1914 GATCTCTGATGATCTTAAATTTTCTTGAAGCGGTCCGAAGATGGGATATCTCTGTC 1973  
DB 315 CATCTTTTAGAAGACATCAAGCTTGTTCACAGAAGCAGTTAAAGATGGGATCTGTCTGT 374  
QY 1974 TGATAGCGTTCGGGATATCAAGTTGCTTGAAGTTTCTTAGGCTGTGTACACACAGTGA 2033  
DB 375 TAGACGAATGCCAG--AACATCAAAACATGCTACCTGGGTTGTACAACTTGTGA 431  
QY 2034 ATGGAATTTGAAAAGATGGACTCAAGGAACAAGCCCGTGTATGCTGGGGTATCTTCGAA 2093  
DB 432 ATGAATAGCAGAAGAGGGCGCAAGAGACAGGGTGTGACGTACTAGGCTACATTAGAA 491  
QY 2094 AAGTATGGGAGGGCTTGCTCGCATCGTATACCAAGAGCCGAATGCTCGCGCAGCAAGT 2153

```

Db      492  ATTTATGGAGATCTGGTTGAACATATACGAAGAGCAGATGTCGGAGCTGCTT 551
      2154  ATGCGCGACATTCAACGAATATGTGAAA 2183
      552  ATGCGCCCTCTCTCAATGAATACATAGAGA 581

RESULT 14
BU971902
LOCUS   BU971902
DEFINITION   HB20A06r BC Hordeum vulgare subsp. vulgare cDNA clone HB20A06
5-PRIME, mRNA sequence.
ACCESSION   BU971902
VERSION     BU971902.1 GI:24222695
SOURCE      EST.
ORGANISM    Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Radchuk, V., Zhang, H., Weschke, W., Potokina, E. and Wobus, U.
Barley ESTs from developing seeds
Unpublished (2002)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 20 row: A column: 6
Seq primer: M13rev.
Location/Qualifiers
1..624
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:236738"
/db_xref="taxon:112509"
/clone="HB20A06"
/tissue_type="developing caryopsis"
/dev_stage="8-15 DAP (days after pollination)"
/lab_host="X110-Gold"
/clone_lib="BC"
/note="vector: pBluescript SK+; Site_1: EcoRI (5'-end of
cDNA); Site_2: XhoI (3'-end of cDNA); developing
caryopsis, 8-15 DAP(days after pollination) Due to a
cloning artefact caused by the kit, in most cases the
EcoRI site is NOT present, as well as the EcoRI adapter
used for cloning. To excise the insert, restriction sites
upstream EcoRI should be used (e.g. BamHI, SalI, PstI).
NOTE: Also due to the cloning system used Blue/White
selection for recombinants is not 100% reliable."

ORIGIN
Query Match 6.8%; Score 184.4; DB 13; Length 624;
Best Local Similarity 59.5%; Pred. No. 2e-22;
Matches 368; Conservative 0; Mismatches 241; Indels 9; Gaps 3;

Qy      1082  GATCTATTGAACCCCTCTGGCGTGTGATACAGTTGAAGCGTTGGGAATCGATCGCTAT 1141
      11  GATCTCTTTGAGCGCATCTGGTCGTCGATCGGTTAGCGCTTGGAAATCTCGCGTAT 70

Qy      1142  TTCAAGAAAGAAATCAAGAAATCTCTCGATTAAGTTATAGGTACTGGGACGCCGAAAGA 1201
      71  TTCAAGCAAGAAATCAAGCAAGTGTGGATTATGTTACAGGCACTGG---ACTGATGAG 127
      1202  GCGTGGGATGGGCAAGATGCATCTTCTCTGATGTCGATGACATGCGCATGGTCTT 1261

```

```

Db      128  GGGATTTGCTGGCGAGGAATCCACTGTAATAGACGTGGATGATACATCCATCCGCAATC 187
      1262  AGAATCTCTGAGACTTCAATGATACAATGATATCTTCAATGATTTCTGGAGAAATTTGAGAGAC 1321
      188  CGGCTGCTGGGCTTCAATGATAGATGTCCTCCCAAGTGTATTTGGGAAGTT---TSGAG 244
      1322  GAGAAAGGAGACTTCTTTGCTTTGCGGTCGCAACGCAAAATTTGGTGTGACCGATAATCTT 1381
      245  AAGGACGGGAGTTCTTCTGTTCTGGGGCAATCAACACAGCAGTCACTGGGATGTAC 304
      1382  AACCTTTATAGATGTTTCAAGATGATGTTTTCGCGGAGAAAGATAATGGAAGAGCTTAAG 1441
      305  AACCTGAACAGGCGCTCTCAGGTAAGTTCCTCGGAGAGGACTTGTGTCAGCGGTGCGAGG 364
      1442  ACCTTCACTACAAATCATCTCCAAATGCTCTTGGCCAAAACCAACGCAATTTGATAAGTGG 1501
      365  AGATTCTCGTATGAGTTCTTATAGAAAAGGGAAGCCCATGGCAGATTCGAGACAAATGG 424
      1502  GCTGTCAAGAAAGCATCTTCTCGGAGAGGTGAGTATGCTATAAGTATCCGTGSCATAGA 1561
      425  ATCAITGCTAAGGATCTTACCAGCGGAGGTAAATATACACTGGACTTCCCATGATGTGA 484
      1562  AGTATGCCAAGATTTGGAGGCAAGAAGTTACATAGAGCAATTT---GGATCAAAATGATGTC 1618
      485  AGCTTACCGGCTGTAGAGCAAGAGCTACTAGATCAATATGCGGTGATATGATGTC 544
      1619  TGCTGGGGAAGACTGTGTATAAGATGCTATATGTGAGCAACGAAATAATTTGGAGCTG 1678
      545  TGGATTGGGAAGACACTCTACAGGATGCTACTTGTGAACAACAACACCTATCTTGAGTTG 604
      1679  GCCAAATTGGACTTCAAT 1696
      605  GCAAAAGCGTGAATTCAT 622

RESULT 15
CA223665
LOCUS   CA223665
DEFINITION   SCJFFL1C04F02.g FL1 Saccharum officinarum cDNA clone SCJFFL1C04F02
5', mRNA sequence.
ACCESSION   CA223665
VERSION     CA223665.1 GI:35280237
KEYWORDS    EST.
SOURCE      Saccharum officinarum
ORGANISM    Saccharum officinarum
REFERENCE   1 (bases 1 to 688)
AUTHORS    Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
TITLE      The libraries that made SUCEST
JOURNAL    Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT    Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: C04 row: F column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..688
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFFL1C04F02"
/lab_host="DH10B"
/clone_lib="FL1"
/note="Organ: Inflorescence at beginning of development"

```

(1cm-long); Vector: pSport1; Site\_1: SalI; Site\_2: NotI;  
An unidirectional cDNA library generated from  
(inflorescence at beginning of development (1cm-long)).  
cDNA was prepared from polyA+ mRNA using SuperScript  
Plasmid System Kit (Invitrogen). The double-strand cDNAs  
were fractionated in a sepharose CL-2B 40cm-columns and  
fragments sizing between 0.8 and 1.5 Kb were  
directionally cloned into the vector. Details of each  
source of RNA and library construction can be obtained at  
<http://sucsest.lad.ic.unicamp.br/public>

ORIGIN

Query Match	6.8%;	Score 184.2;	DB 14;	Length 688;
Best Local Similarity	57.1%;	Pred. No. 2.1e-22;		
Matches 396;	Conservative 0;	Mismatches 288;	Indels 9;	Gaps 3;
QY 1154	ATCAAGAAATCTCTGGATTACGTTTATAGTACTGGACGCCGAAAGAGCGGTGGGATGG	1213		
Db 2	ATTAAGCAATGATGGACTATGTGAACAGGCACTGG---ACTGAAGATGGGATTTGCTGG	58		
QY 1214	GCAAGATGCAATCCATTCTCTGATGTCGATGACACTGCCATGGGTCTTAGAATCCTGAGA	1273		
Db 59	GCTAGGACTCTACTGTAAGATGTGGATGACAGCTATGGCTTCCGACTACTAAGG	118		
QY 1274	CTTCATGGATCAATCTATCTTCAGATGTTCTGGAGATTTTCAGAGACGAGAAAGAGAC	1333		
Db 119	CTACATGGATCAATGTCTCCCAAGTGTGTTAGAAATTT---TGAGAAAGATGGAGAG	175		
QY 1334	TTCTTTTGGCTTTGGCGGTCAAAACGCAAAATGGTGTGACCGATAATCTTAAACCTTTATAGA	1393		
Db 176	TTCTTTTGGTGTGGGCAATCACTCAAGCGCTCACTGGGATGTACAACCTCAACAGA	235		
QY 1394	TGTTCAAGATGTTTTCGGGAGAAAGATATGGAAGAGCTAAGACCTTCACTACA	1453		
Db 236	GCCTCTCAAAATAGCTTTCAAGGGGAGGATGATTACACCGTGTAGGACTTTCTCATAT	295		
QY 1454	AATCATCTCCAAATGCTCTTGCCAAAACAAACGCAATTTGATAAGTGGGCTGTCAAGAGAG	1513		
Db 296	GAGTTTCTCAGACAAAGAGAGCCCAAGGATGCTCCGTGATATAATGGATCATCGCGNAG	355		
QY 1514	GATCTTCTCGAGAGTGGAGTATGCTATAAGTATCCGTGGCATAGAAGTATGCCAAGA	1573		
Db 356	GATCTAGCTGGCGAGGTACAATATACACTAGACTTCCCTTGGTATGCAAGCTTGCCTCGT	415		
QY 1574	TTGGAGGCAAGATTTACATAGACCAATTTGGATCAA---ATGATCTCTGGCTGGGGAAG	1630		
Db 416	GTGAGGGCAAGAACCTTATCTAGATCAGTATGGCGGTAAAGATGATGTTGGATTGGAAAG	475		
QY 1631	ACTGTGTATAAGATGCTATATGTGAGCAACGAAAAATATTGGAGCTGGCCAAATTGGAC	1690		
Db 476	ACCTTATACAGGATGCTCTTGTGAATAACGACATATATCTTGATTTGGCAATAATGGAT	535		
QY 1691	TTCAATATGGTGGAGGCTTTACACCAAGGAGACTCAACACATTGTCTAGCTGGTGGAGA	1750		
Db 536	TTCAACCGTTGCCAAGCTCTACATCAGCTTGTAGTGTAAAGAGCTTCAAAATGTGGTACATA	595		
QY 1751	GAATCGGATTCATGATCTTACATTCACCCCGCAGCGGCTGTGGAAATGTATTCTCA	1810		
Db 596	GAGAAATTGCCTTGACACTTTTGGAGTGCACCGCAAGATGTTTTAAGAGCTTATTTTAA	655		
QY 1811	GTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCT	1843		
Db 656	GCTGGGCTTTGCAATTTATGAACCTTAGTCTGCT	688		

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:53:08 ; Search time 54.4614 Seconds  
(without alignments)  
4529.153 Million cell updates/sec

Title: US-10-041-007-2

Perfect score: 4645

Sequence: 1 MAGVLPANLPCSLQLSPKP.....GISDKEMKHVRLTDFPVA 873

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4645	100.0	873	ABU09779	Maidenhair
2	2844	61.2	868	AAy90858	Grand fir
3	2844	61.2	868	ABB79393	Taxadiene
4	2825	60.8	868	ABU09780	White fir
5	2114.5	45.5	817	AAy06567	E-alpha-b
6	2114.5	45.5	817	AAy06562	Grand fir
7	2114.5	45.5	817	ABU09781	White fir
8	2113.5	45.5	817	AAy06566	Grand fir
9	2113.5	45.5	817	AAy06568	E-alpha-b
10	2113.5	45.5	817	ABB79391	Taxadiene
11	2085.5	44.9	782	AAW85703	Grand fir
12	2085.5	44.9	782	AAy90853	Grand fir
13	2085.5	44.9	782	AAW85703	Grand fir
14	2085.5	44.9	782	ABB79392	Taxadiene
15	1959	42.2	862	AAW31655	Pacific y
16	1959	42.2	862	AAy90852	Yew taxad
17	1959	42.2	862	AAW79394	Taxadiene
18	1959	42.2	862	ABU09782	Pacific y
19	1377.5	29.7	815	AAE24916	Soybean c
20	1377.5	29.7	815	AAE24916	Soybean c
21	1332	28.7	802	AAy90848	Gibberell
22	1332	28.7	802	ABB92819	Herbicida
23	1317	28.4	823	AAy93800	Anthrac ea
24	1255	27.0	581	AAy06569	Delta-sel
25	1252	27.0	580	AAy06571	Delta-sel

26	1252	27.0	581	2	AAy06570	Delta-sel
27	1252	27.0	581	2	AAy06563	Grand fir
28	1252	27.0	581	3	AAy90854	Grand fir
29	1215.5	26.2	630	4	AAW69391	Grand fir
30	1206	26.0	577	2	AAW85704	Grand fir
31	1206	26.0	577	4	AAW69374	Grand fir
32	1194.5	25.7	637	2	AAW85702	Limonene
33	1194.5	25.7	637	3	AAy90859	Grand fir
34	1194.5	25.7	637	4	AAW69372	Grand fir
35	1193	25.7	637	4	AAW69392	Grand fir
36	1173.5	25.3	627	2	AAW85710	Grand fir
37	1173.5	25.3	627	4	AAW69380	Grand fir
38	1173.5	25.3	627	5	ABB79395	Taxadiene
39	1167.5	25.1	627	2	AAW85700	Myrcene s
40	1167.5	25.1	627	3	AAy90842	Grand fir
41	1167.5	25.1	627	4	AAW69370	Grand fir
42	1151.5	24.8	593	2	AAy06573	Gamma-hum
43	1151.5	24.8	593	2	AAy06572	Gamma-hum
44	1151.5	24.8	593	2	AAy06564	Grand fir
45	1151.5	24.8	593	3	AAy90855	Grand fir

## ALIGNMENTS

## RESULT 1

ABU09779

ID ABU09779 standard; protein; 873 AA.

AC ABU09779;

DT 06-AUG-2003 (first entry)

DE Maidenhair tree levopimaradiene synthase.

KW Levopimaradiene synthase; ginkgolide; enzyme; Escherichia coli;

KW Candida albicans; Kluyveromyces lactis; Saccharomyces; maidenhair tree.

OS Ginkgo biloba.

PN US2002164736-A1.

XX 07-NOV-2002.

XX 07-JAN-2002; 2002US-00041007.

XX 05-JAN-2001; 2001US-0259881P.

XX (MATS/) MATSUDA S P T.

XX (SCHE/) SCHEPMANN H G.

XX Matsuda SPT, Schepmann HG;

XX WPI; 2003-352584/33.

XX New purified and isolated nucleic acid sequence encoding a levopimaradiene synthase useful for producing pure ginkgolides in cells.

XX Claim 20; Fig 4; 37pp; English.

XX The invention relates to a purified and isolated nucleic acid sequence encoding a levopimaradiene synthase. The nucleic acid is used for the preparation of a ginkgolide in a cell, an expression vector, a unicellular organism (e.g. Saccharomyces, Escherichia coli, Candida albicans or Kluyveromyces lactis), a yeast host cell, a plant (particularly Ginkgo biloba), a host cell, a transgenic plant or a seed of the transgenic plant, for the biosynthesis of ginkgolides and in molecular biology, molecular genetics, and organic chemistry. The nucleic acid is capable of producing large amounts of pure ginkgolides in a cost-effective manner, particularly in an organism capable of a high-yield ginkgolide-producing system. This sequence represents a levopimaradiene synthase polypeptide of the invention

25-AUG-2000	(first entry)
Grand fir abietadiene synthase protein sequence SEQ ID NO:56.	
Synthase; protein co-ordinate data; active site; modification; terpenoid;	
3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;	
isoprenoid; breeding programme; fragrance; flavour; pheromone;	
defensive agent; pigment; antitumour; steroid hormone;	
signal transduction pathway; bile acid; affinity purification;	
photoreceptor; enzymatic synthesis; nutrient supplement;	
immunological reagent.	
Abies grandis.	
WO200017327-A2.	
30-MAR-2000.	
17-SEP-1999; 99WO-US021419.	
18-SEP-1998; 98US-0100993P.	
22-APR-1999; 99US-0130628P.	
23-AUG-1999; 99US-0150262P.	
(KENT ) UNIV KENTUCKY RES DEPT.	
(SALK ) SALK INST BIOLOGICAL STUDIES.	
Chappell J, Manna KR, Noel JP, Starke CM;	
WPI; 2000-292839/25.	
N-PSDB; AAA38937.	
Novel terpene synthase enzymes, useful for producing terpene	
hydrocarbons, e.g. fragrances or antitumor agents, are derived from known	
enzymes by specific amino acid alterations.	
Claim 112; Page 443-445; 450pp; English.	
The present invention describes an isolated terpene synthase (I)	
comprising a region with at least 20% identity to region 265-535 of a 548	
amino acid (aa) sequence (Ia), given in AA90831. (I) contains nine alpha	
-carbon atoms (alphaC) that have interatomic distances, between each	
other, within tabulated ranges, have a centre point (within a sphere of	
radius 2.3 Angstrom) within tabulated ranges, and have an ordered	
arrangement of R groups (defining aa side chains), excluding specific	
tabulated arrangements (tables given in the specification). (I), and	
related enzymes, are used to produce a wide range of terpenoids (e.g.	
cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,	
flavours, pheromones, defensive agents, pigments, antitumour agents,	
components of signal transduction pathways, precursors of steroid	
hormones and bile acids, as photoreceptors, and as co-factor side chains.	
Some synthases with little or no catalytic activity (and nucleic acids	
encoding them) are used as controls in the analysis of products formed	
by enzymatic synthesis; as nutrient supplements; for affinity purification	
of isoprenoids; or to develop immunological reagents or nucleic acids for	
monitoring expression of terpene synthase or inheritance of the gene in	
plant breeding programs. The new synthases may produce novel terpene	
products. AAA38910 to AAA38938 and AA90831 to AA90859 represent	
sequences used in the exemplification of the present invention	
Sequence 868 AA;	
Query Match 61.2%; Score 2844; DB 3; Length 868;	
Best Local Similarity 62.3%; Pred. No. 1.9e-248;	
Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11	
Oy 12 SLOLSPKVP---PRQSTNIIIPFKHSSFGFNAQHCVSHLRWN-----CVGIHA 60	
Dd 5 SSLSISQIPTAAHLHTANAQSIPHSTTLNAGSSAKRSRLYRWGKSNIACVG--E 62	
Oy 61 SAAETRPDQLPQEERFVS---LNADYHPAVWKDDFDLSLTSPNSHATSKSVDETKNR 117	
Dd 63 GGATSPVYQSAEKNDLSLSSTLVVKREFFPGFWKKDLIDSLTS--SHKVAAS--DE---KR 115	



```
Qy 118 IOTLVKEIQCFQSGDGETNPSAYDTAWVARIPISDGSGAPQPPOTLQWLNQLPDGS 177
Db 116 IETLISEIKNFRCMGGETNPSAYDTAWVARIPIAVDGSNPHFPETVEITLQNLQKDG 175
Qy 178 WGEBCIFLAYDRVLNTLACLLTLKINWKGDIQVQKGVFVRKHEEMKDEADNHRPSGFE 237
Dd 176 WGEBCIFLAYDRVLNTLACLLTLKINWKGDIQVQKGVFVRKHEEMKDEADNHRPSGFE 235
Qy 238 VVFPAMLDKAKSGLDLPYHLPTFSIQHQRKQKLOKIPLVNLHNHOTALYSLEGLQDV 297
Dd 236 IVFPAMLDKAKSGLDLPYHLPTFSIQHQRKQKLOKIPLVNLHNHOTALYSLEGLQDV 295
Qy 298 VDMQKITNLSQSDGSLSSPASTACVPMHTONKRCLEHFLNVLKSGDYVPCYPLDLFE 357
Dd 296 VDMQKITNLSQSDGSLSSPASTACVPMHTONKRCLEHFLNVLKSGDYVPCYPLDLFE 355
Qy 358 RLWAVDVTVERLIGIDRYFKKEIKESLDVYVYRWDAERGVGWACNPIPDVDDTANGLRILR 417
Dd 356 RLWAVDVTVERLIGIDRYFKKEIKESLDVYVYRWDAERGVGWACNPIPDVDDTANGLRILR 414
Qy 418 LHGVNVSDDVLENFRDEKGDFFCFAGQTQIGVTDNLNLYRCSQVCFPGKIMEAKTFTT 477
Dd 415 LHGVNVSDDVLENFRDEKGDFFCFAGQTQIGVTDNLNLYRCSQVCFPGKIMEAKTFTT 474
Qy 478 NHLONALAKNNAFDKNAVKKDLPGVEYAIKYPWHRSMRPLEARSYLEQSGNDVWLKGT 537
Dd 475 RYLENALLENVDAPDKWAFKKNIRGEVEYALKYPWHRSMRPLEARSYLEQSGNDVWLKGT 534
Qy 538 VYKMLVYSNEKYLELAKLDFNMVQALQKQETQHTIVSWWRSGGNDLPTQRVPMYPSV 597
Dd 535 VYKMLVYSNEKYLELAKLDFNMVQALQKQETQHTIVSWWRSGGNDLPTQRVPMYPSV 594
Qy 598 AVSMFEPEAFACRIAYARTKSLAVILDDLYDTHGSLDDKLDFSEAVRRWDISVLDSVRDN 657
Dd 595 ASPFEPEFSCREVTYKTSNFTVILDDLYDTHGSLDDKLDFSEAVRRWDISVLDSVRDN 653
Qy 717 QKVCPLGLNTYNGFQDGLKEGRDVLGYLRKVGWGLLASITYKEAWSAAKYVPTFNE 717
Dd 654 QMKICFCFGYNTFNDIAKEGRERQGRDVLGYLRKVGWGLLASITYKEAWSAAKYVPTFNE 713
Qy 718 YVENAKYSIALATVVLNSIPTGELLPDYILQVVDLRSKFLHLVSLGRINDTKYQAE 777
Dd 714 YVENAKYSIALATVVLNSIPTGELLPDYILQVVDLRSKFLHLVSLGRINDTKYQAE 773
Qy 778 RNRGELVSSVQCYNRENPECTEERBEALSHVYGIIDNALKEMLWELANPASNAPLCVRELLF 837
Dd 774 RNRGELVSSVQCYNRENPECTEERBEALSHVYGIIDNALKEMLWELANPASNAPLCVRELLF 831
Qy 838 NTARVMOLFMYNRDGFGIS-DKEMKDHVSRTLFDPVA 873
Dd 832 ETARIMOLFMYNRDGFGIS-DKEMKDHVSRTLFDPVA 868
RESULT 3
ABB79393
ID ABB79393 standard; protein; 868 AA.
XX AC ABB79393;
XX AC ABB79393;
XX DT 13-AUG-2002 (first entry)
XX DE Taxadiene synthase homologous protein sequence q38710.
XX KW Taxane; taxane synthesis; taxadiene synthase; cytostatic; anticancer;
XX KW enzyme.
XX OS Unidentified.
XX FN WO2002040694-A2.
XX PD 23-MAY-2002.
XX
```

```
PF 16-NOV-2001; 2001WO-DK000763.
XX 17-NOV-2000; 2000DK-00001730.
PR 29-NOV-2000; 2000US-0253843P.
XX PA (NOVO ) NOVOZYMES AS.
XX PI Vind J;
XX WPI; 2002-471624/50.
DR Heterologous production of taxane analogs related compound useful as
XX anticancer agent involves cloning a full-length taxane synthesis pathway
PT from a taxan-producing organism into a taxane-resistant host cell.
PT Example 3; Fig 1; 60pp; English.
PS The present invention describes the heterologous expression of taxane
XX analogues, comprising cloning a DNA sequence comprising a taxane
CC synthesis pathway, making a DNA construct in which the DNA sequence is
CC under control of regulatory elements, introducing the DNA construct into
CC a host cell, growing the host cell to produce taxane analogues, and
CC recovering the required taxane from the culture medium. Also described:
CC (1) an isolated DNA sequence comprising the taxane synthesis pathway;
CC (preferably taxol synthesis pathway); (2) an expression vector comprising
CC the isolated DNA sequence; and (3) a host cell comprising a taxane
CC synthesis pathway derived from a taxane-producing organism, plant or
CC tree, the taxane synthesis pathway is foreign to the host cell. The
CC taxane analogues have cytostatic activity, and so can be used as
CC anticancer agents. The present sequence represents a taxadiene synthase
CC homologous protein, which is given in an example from the present
XX invention
XX SQ Sequence 868 AA;
Query Match 61.2%; Score 2844; DB 5: Length 868;
Best Local Similarity 62.3%; Pred. No. 1.9e-248; Indels 28; Gaps 11;
Matches 546; Conservative 120; Mismatches 183;
Qy 12 SLQLSPKVP---FRQSTNIIIPFKKSSFGFNACQHCVRSHLRWN-----CVGIHA 60
Dd 5 SSSLSQIPTAHHLTANAQSIPIHFSITTLNAGSASKRSILYLRWGSKNKIACVG--E 62
Qy 61 SAETRPDQLPQBERFVSR---LNADYHPAPWKKDFIDSLTSPNSHATSSVDDETINKR 117
Dd 63 GGATSVPYQSAEKNDLSSTLTKVREPFGFWKDDLSLTS--SHKVAAS--DE--KR 115
Qy 118 IOTLVKEIQCFQSGDGETNPSAYDTAWVARIPISDGSGAPQPPOTLQWLNQLPDGS 177
Dd 116 IETLISEIKNFRCMGGETNPSAYDTAWVARIPIAVDGSNPHFPETVEITLQNLQKDG 175
Qy 178 WGEBCIFLAYDRVLNTLACLLTLKINWKGDIQVQKGVFVRKHEEMKDEADNHRPSGFE 237
Dd 176 WGEBCIFLAYDRVLNTLACLLTLKINWKGDIQVQKGVFVRKHEEMKDEADNHRPSGFE 235
Qy 238 VVFPAMLDKAKSGLDLPYHLPTFSIQHQRKQKLOKIPLVNLHNHOTALYSLEGLQDV 297
Dd 236 IVFPAMLDKAKSGLDLPYHLPTFSIQHQRKQKLOKIPLVNLHNHOTALYSLEGLQDV 295
Qy 298 VDMQKITNLSQSDGSLSSPASTACVPMHTONKRCLEHFLNVLKSGDYVPCYPLDLFE 357
Dd 296 VDMQKITNLSQSDGSLSSPASTACVPMHTONKRCLEHFLNVLKSGDYVPCYPLDLFE 355
Qy 358 RLWAVDVTVERLIGIDRYFKKEIKESLDVYVYRWDAERGVGWACNPIPDVDDTANGLRILR 417
Dd 356 RLWAVDVTVERLIGIDRYFKKEIKESLDVYVYRWDAERGVGWACNPIPDVDDTANGLRILR 414
Qy 418 LHGVNVSDDVLENFRDEKGDFFCFAGQTQIGVTDNLNLYRCSQVCFPGKIMEAKTFTT 477
Dd 415 LHGVNVSDDVLENFRDEKGDFFCFAGQTQIGVTDNLNLYRCSQVCFPGKIMEAKTFTT 474
Qy 478 NHLONALAKNNAFDKNAVKKDLPGVEYAIKYPWHRSMRPLEARSYLEQSGNDVWLKGT 537
Dd 475 RYLENALLENVDAPDKWAFKKNIRGEVEYALKYPWHRSMRPLEARSYLEQSGNDVWLKGT 534
```

475 RYLRLNLENVDAFDKWAFFKQNGIRGEVEYALKYPMHKSMPRLSEARSYIENYGPDDVWLKGT 534  
538 VYKMLYVNEKYLELAKLDFNMVQALHQKQTHIVSWWRSGFNDLFTTRORPVMEMFVS 597  
535 VYMPFYISNEKYLELAKLDFNMVQALHQKQTHIVSWWRSGFNDLFTTRORPVMEMFVS 594  
598 AVSMPEPEPAACRIAYAKTSCLAVIDLDLYTHGSLDLDLKLFESEAVRWDSVLDSVRDN 657  
595 ASFIPEPEFSKREYVTKTSNFTVILDDLYDAHGSLLDLDLKLFESEAVRWDSVLDSVRDN 653  
658 OLKVCFLGLYNTVNGFGKDGKLEQGRDVLGRLKRWEGLLASYTEAEWSAAKYVPTTNE 717  
654 QMKICFVGFTNTFNDIAKEGRQRGRDVLGRLKRWEGLLASYTEAEWSAAKYVPTTNE 713  
718 YVENAKVSIALATVVLNSIFPTGELLDPYILQVVDLRSKFLHLVSLTGRLLNDTKTYQ 777  
714 YIENASVSIALGTVVLISALFTGEVLTDEVLSDKIDRESRFLQMLGLTGRVNDTKTYQ 773  
778 RNRGELVSVQCYMRENPCTEERALSHTVYGIIDNALKELNWLNPANPLCVRRLIF 837  
774 RQGEVAGASAIQCYMKDHPKISEEALQHVYSVMENALEELNREFVN--NKIPDIYKRLVF 831  
838 NTAHVQMLFYMYRDRDGFIS--DKEMKDHVSRTLPDPVA 873  
832 ETARIMQLFYMQGDGLTSLHDMWEIKHVKNCFLQFPA 868

RESULT 4  
ID ABU09780  
XX AC ABU09780;  
XX DT 06-AUG-2003 (first entry)  
XX DE White fir abietadiene synthase.  
XX KW Levopimaradiene synthase; ginkgolide; enzyme; Escherichia coli;  
XX KW Candida albicans; Kluyveromyces fragilis; Saccharomyces; white fir;  
XX KW abietadiene synthase.  
XX OS Abies grandis.  
XX PN US2002164736-A1.  
XX PD 07-NOV-2002.  
XX PF 07-JAN-2002; 2002US-00041007.  
XX PR 05-JAN-2001; 2001US-0259881P.  
XX PA (MATSU) MATSUDA S P T.  
XX PA (SCHE) SCHEPMANN H G.  
XX PI Matsuda SPT, Schepmann HG;  
XX WI; 2003-352584/33.  
XX PT New purified and isolated nucleic acid sequence encoding a  
XX PT levopimaradiene synthase useful for producing pure ginkgolides in cells.  
XX PS Example 3; Fig 4; 37pp; English.

The invention relates to a purified and isolated nucleic acid sequence encoding a levopimaradiene synthase. The nucleic acid is used for the preparation of a ginkgolide in a cell, an expression vector, a unicellular organism (e.g. Saccharomyces, Escherichia coli, Candida albicans or Kluyveromyces fragilis), a yeast host cell, a plant (particularly Ginkgo biloba), a host cell, a transgenic plant or a seed of the transgenic plant, for the biosynthesis of ginkgolides and in molecular biology, molecular genetics, and organic chemistry. The nucleic acid is capable of producing large amounts of pure ginkgolides in a cost-effective manner, particularly in an organism capable of a high-yield

ginkgolide-producing system. This sequence represents an abietadiene synthase polypeptide which shows similarity to the levopimaradiene synthase of the invention

Query Match  
Best Local Similarity 60.8%; Score 2825; DB 6; Length 868;  
Matches 543; Conservative 124; Mismatches 180; Indels 32; Gaps 12;  
XX Sequence 868 AA;

12 SLQSPKVP-----FRQSTNLIIPFKHSSGFNAQHCVRSHLRN-----CVGI 58  
5 SSSLSQIPTAHHLTANTQSPHSTTLNAGSSASK--RRSLYLKRWGKGNKIACVG- 61  
59 HASAAETRPDOLPOBEREVSF---LNADYHPAVMKDDFIDSLTSPNSHATSKSSVDETIN 115  
62 -EGGATVPYQSAEKNDLSSTLVKREFPGFWKDDLDLSLT--SHKVAAS--DE--- 113  
116 KRIOTLVKEIOCMFQSGMDGETNPSAYDTAWARIPTSDGSGAPQFOTLOWILNNQLPD 175  
114 KRIETLISEIKNFRCMGGETNPSAYDTAWARIPTSDGSGAPQFOTLOWILNNQLPD 173  
176 GSWGEECIFLAYDRVLNLTACLLTLKWNKGDIOVQKGVFVRKHMEEMKDEADNHRPSG 235  
174 GSWGEGFYFLAYDRILATLACIITLWRTGETVQKGIESTRTQAGKMEDEADNHRPSG 233  
236 FEVFPAPMLDEAKSLGLDPYHLPFISQIHOKRQKLOKIPUNLVHNNHTALLYSLEGLQ 295  
234 FEIVFPAMLEAKSLGLDPYDLPLKQIIEKREAKRIPDTDLVYALPTTLLYSLEGLQ 293  
296 DVVDQWETNLOSROSGFLSSPASTACVPMHTONKRLHFLNVLKSFQDGYVCHYPLDL 355  
294 EIVEWEKIMKLSQSGFLSSPASTAAVPMRTGNKCKLDLFLNVLKSGHNVCHYPLDL 353  
356 PERLWAVDTVERLIGIDRYFKKEIKESLDYVRYWDAERGVGWARGNPIPDVDDTAMGLRI 415  
354 PERLWAVDTVERLIGIDRYFKKEIKESLDYVRYWDAERGVGWARGNPIPDVDDTAMGLRI 412  
416 LRLHGVNVSVDYLENFRDEKGFPCFAGTQIGVTDNMLYRCQVCFPGKIMEAKTF 475  
413 LRLHGVNVSVDYLENFRDEKGFPCFAGTQIGVTDNMLYRCQVCFPGKIMEAKTF 472  
476 TTNHLONALAKNNAFKWAKKDLGPEVEYALKYPMHRSMPRLSEARSYTEQFSGNDVWL 535  
473 TERYLRNALLENVDADFKAFFKNGIRGEVEYALKYPMHKSMPRLSEARSYTEQFSGNDVWL 532  
536 KTVYKMLYVNEKYLELAKLDFNMVQALHQKQTHIVSWWRSGFNDLFTTRORPVMEMF 595  
533 KTVYKMLYVNEKYLELAKLDFNMVQALHQKQTHIVSWWRSGFNDLFTTRORPVMEMF 592  
596 SVAVSMPEPEPAACRIAYAKTSCLAVIDLDLYTHGSLDLDLKLFESEAVRWDSVLDSVR 655  
593 SPASFIFEPEFSKREYVTKTSNFTVILDDLYDAHGSLLDLDLKLFESEAVRWDSVLDSVR 651  
656 DNQKVCFLGLYNTVNGFGKDGKLEQGRDVLGRLKRWEGLLASYTEAEWSAAKYVPTT 715  
652 PKQMKICFVGFTNTFNDIAKEGRQRGRDVLGRLKRWEGLLASYTEAEWSAAKYVPTT 711  
716 NEXVENAKVSIALATVVLNSIFPTGELLDPYILQVVDLRSKFLHLVSLTGRLLNDTKTYQ 775  
712 NEYIENASVSIALGTVVLISALFTGEVLTDEVLSDKIDRESRFLQMLGLTGRVNDTKTYQ 771  
776 AERNRCELVSQCYMRENPCTEERALSHTVYGIIDNALKELNWLNPANPLCVRRL 835  
772 AERQGEVAGASAIQCYMKDHPKISEEALQHVYSVMENALEELNREFVN--NKIPDIYKRL 829  
836 LFNARVQMLFYMYRDRDGFIS--DKEMKDHVSRTLPDPVA 873  
830 VFETARIMQLFYMQGDGLTSLHDMWEIKHVKNCFLQFPA 868

RESULT 5  
AAV06567

ID	AA06567	standard; protein; 817 AA.	QY	254	LPYHLPTISQIHOKRQKLOKIPLNVLHNNHOTALLYSLLEGLODQVVDWQSEITNLSQDQSP	313
XX						
AC	AA06567;		Db	197	LPYDLPTIHLQTKRQERLAKLSRETYAVPSPLYSLEGIQDIWEVERIMEVQSDQSP	256
XX						
DT	08-OCT-1999	(first entry)	QY	314	LSSPASTACVFMHTONKRCCLHFLNVLKSGDYVPCYVPLDPLPERLWAVDVTVRGLIDRY	373
XX						
DE	E-alpha-bisabolene	synthase computer-generated sequence.	Db	257	LSSPASTACVFMHTGDAKCLEFLNSVMKFGNVPCLYPVDLLERLLIDVNIYRGLIYR	316
XX						
KW	E-alpha-bisabolene	synthase; grand fir; sesquiterpene synthase;	QY	374	PKBIKESLDYVRYDAERGVMGWARCNPIPDVDDTAMGURLRLHGVNVSDDLNPDR	433
KW	gymnosperm; sesquiterpenoid;	disease resistance; fragrance; flavour.	Db	317	PEKEIKEALDYVYRHMN-ERGIGWGRNLPIADLETALGFRLRLHRYNVSIPAIFNPKD	375
XX						
OS	Synthetic.		QY	434	EKGDFPCFAGOTQIGVTDNLNLVRCOVCPGEGKIMESAKTFTTNHLQNALAKNAPDKW	493
XX	Abies grandis.					
XX			Db	376	ANGKPCISTGQFNKDVASMLNLYASQALPFGENILDEAKSFATKYLREALEKSETSSAW	435
FH	Key	Location/Qualifiers	QY	494	AVKXDLPGVEVEYAIKYPWHRSMPLREARSYTEQFGSNDVWLGKTVYKMLYVSNKYLELA	553
FT	Peptide	/note= "conserved element, Asp residues function in				
FT		coordination of a divalent cation"	Db	436	NNQNLSQEIYKALKTSWHAHVPRVAKRYCQVYRDPYARIAKCVYKLPYVNNKFLGLG	495
XX						
PN	WO9937139-A1.		QY	554	KLDPNMOALHQKETQIHIVSWRSGFNDLTFTTQRPVEMVYFSAVSMFEPFAACRIAY	613
XX						
PD	29-JUL-1999.		Db	496	KLDENIIQSIHQEEMKNVTWFRDGLPLFTFARERPLEFYFLVAAGTYEPQYAKRPLF	555
XX						
XX	21-JAN-1999;	99WO-US001300.	QY	614	AKTSLAVILDDLVDYTHGSLDDLKLPSEAVRRWDISVLDSVRDNQLKVCFLGLYNTWNGF	673
XX						
PR	22-JAN-1998;	98US-0072204P.	Db	556	TKVACLQTVLDDMYDTYGTLDLDELKLFTEAVRRWDLSTFENLPD-YMKLCYQIYYDIVEH	614
XX		(UNIW ) UNIV WASHINGTON STATE RES FOUND.				
XX			QY	674	GKGLKEQGRDVLGKRWEGLLASVTKEAENSAKYVPTFNEYVENAKYSIALATVVL	733
PI	Croteau RB, Crock JB, Bohlman J, Jetter R, Steele CL;		Db	615	AWAEKEQGRDELVSFFRKGWEDYLLGYEEAEWLAAYVPTLDYIKNGITSIQRIILL	674
DR	WPI, 1999-458569/38.					
XX	N-PSDB; AAX87530.		QY	734	NSI-FFTGELLPDYILOQVDL--RSKFLHLVSLTGRINDTKTYOABERNRGELVSSVQCY	790
XX			Db	675	SGVLMDGQLLSQALEKVDYVPGRRVLTSLNSLRDLADDTKTYKAEKARGELASSIBCY	734
PT	Grand fir (Abies grandis) sesquiterpene synthase genes and protein		QY	791	MRENPECTEEREAALSHVGIIDNALKEINWELANPASNAPLCVRLLFTFARVMOQLFYMYR	850
PT	products, useful for sesquiterpenoid production.		Db	735	MKDHPCTEEREAALSHVGIIDNALKEINWELANPASNAPLCVRLLFTFARVMOQLFYMYR	850
XX						
PS	Example 11; Page 98-101; 136pp; English.		QY	851	DGFGISDKEMKDHVSRTLFDPV 872	
XX			Db	794	DGFGVSKLEVKDHIKECLIEPL 815	
CC	This is the sequence of an E-alpha-bisabolene synthase polypeptide					
CC	encoded by a computer-generated nucleic acid (see AAX87530) based on an					
CC	isolated grand fir (Abies grandis) E-alpha-bisabolene synthase clone (see					
CC	AAX87504). It contains conservative amino acid substitutions relative to					
CC	the native sequence (see AAY06562). E-alpha-bisabolene synthase is a					
CC	wound-inducible enzyme capable of generating multiple sesquiterpenes,					
CC	especially E-alpha-bisabolene, from farnesyl diphosphate. The invention					
CC	provides gymnosperm sesquiterpene synthase nucleic acids and recombinant					
CC	proteins, as well as expression vectors, and a method of imparting or					
CC	enhancing production of a gymnosperm sesquiterpene synthase in a					
CC	prokaryotic or eukaryotic host cell, especially in a plant, in order to					
CC	impart, enhance or modify production of sesquiterpenoids, and hence to					
CC	alter the flavour or fragrance of plant products, to improve disease					
CC	resistance or to alter ecological interactions mediated by farnesyl					
CC	diphosphate and its derivatives					
XX						
SQ	Sequence 817 AA;					
	Query Match	45.5%; Score 2114.5; DB 2; Length 817;				
	Best Local Similarity	49.5%; Pred. No. 3e-182;				
	Matches	397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;				
QY	79 RLNDYHPAVKDDPIDSITSPNSHATSKSVDETINKRIQTLVKEIQMFQS--MGDGE	136				
Db	25 RRTANPHNPVGVYDLVHLSKSP-----YIDSSYRERAEVLVSEIKAMLNPAITGDGE	76				
QY	137 T--NPSAYDTAWVARIPISDGSGAPQFPQTLOWILNNQLPDGSMGEEICIFLAYDRVNTL	194				
Db	77 SMITPSAYDTAWVARIPISDGSGAPQFPQTLOWILNNQLPDGSMGEEICIFLAYDRVNTL	136				
QY	195 ACLTLTKIWNKGDIGIQVQKGVFVRKHMEEMKDEAD-NHRPSGFEVFPVFPAMDLEAKSLGLD	253				
Db	137 SCVLVLLKWNVGDQLVEQGEIFIKSNLELVKDETDQDLSVTDFFIIFPSILLREAGSLRLG	196				

254 LPYHLPTISQIHOKRQKLOKIPLNVLHNNHOTALLYSLLEGLODQVVDWQSEITNLSQDQSP 313  
 197 LPYDLPTIHLQTKRQERLAKLSRETYAVPSPLYSLEGIQDIWEVERIMEVQSDQSP 256  
 314 LSSPASTACVFMHTONKRCCLHFLNVLKSGDYVPCYVPLDPLPERLWAVDVTVRGLIDRY 373  
 257 LSSPASTACVFMHTGDAKCLEFLNSVMKFGNVPCLYPVDLLERLLIDVNIYRGLIYR 316  
 374 PKBIKESLDYVRYDAERGVMGWARCNPIPDVDDTAMGURLRLHGVNVSDDLNPDR 433  
 317 PEKEIKEALDYVYRHMN-ERGIGWGRNLPIADLETALGFRLRLHRYNVSIPAIFNPKD 375  
 434 EKGDFPCFAGOTQIGVTDNLNLVRCOVCPGEGKIMESAKTFTTNHLQNALAKNAPDKW 493  
 376 ANGKPCISTGQFNKDVASMLNLYASQALPFGENILDEAKSFATKYLREALEKSETSSAW 435  
 494 AVKXDLPGVEVEYAIKYPWHRSMPLREARSYTEQFGSNDVWLGKTVYKMLYVSNKYLELA 553  
 436 NNQNLSQEIYKALKTSWHAHVPRVAKRYCQVYRDPYARIAKCVYKLPYVNNKFLGLG 495  
 554 KLDPNMOALHQKETQIHIVSWRSGFNDLTFTTQRPVEMVYFSAVSMFEPFAACRIAY 613  
 496 KLDENIIQSIHQEEMKNVTWFRDGLPLFTFARERPLEFYFLVAAGTYEPQYAKRPLF 555  
 614 AKTSLAVILDDLVDYTHGSLDDLKLPSEAVRRWDISVLDSVRDNQLKVCFLGLYNTWNGF 673  
 556 TKVACLQTVLDDMYDTYGTLDLDELKLFTEAVRRWDLSTFENLPD-YMKLCYQIYYDIVEH 614  
 674 GKGLKEQGRDVLGKRWEGLLASVTKEAENSAKYVPTFNEYVENAKYSIALATVVL 733  
 615 AWEAEKEQGRDELVSFFRKGWEDYLLGYEEAEWLAAYVPTLDYIKNGITSIQRIILL 674  
 734 NSI-FFTGELLPDYILOQVDL--RSKFLHLVSLTGRINDTKTYOABERNRGELVSSVQCY 790  
 675 SGVLMDGQLLSQALEKVDYVPGRRVLTSLNSLRDLADDTKTYKAEKARGELASSIBCY 734  
 791 MRENPECTEEREAALSHVGIIDNALKEINWELANPASNAPLCVRLLFTFARVMOQLFYMYR 850  
 735 MKDHPCTEEREAALSHVGIIDNALKEINWELANPASNAPLCVRLLFTFARVMOQLFYMYR 850  
 851 DGFGISDKEMKDHVSRTLFDPV 872  
 794 DGFGVSKLEVKDHIKECLIEPL 815  
 RESULT 6  
 AAY06562  
 ID AAY06562 standard; protein; 817 AA.  
 AC AAY06562;  
 XX AAY06562;  
 DT 08-OCT-1999 (first entry)  
 XX Grand fir E-alpha-bisabolene synthase.  
 XX E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;  
 KW gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.  
 XX Abies grandis.  
 FH Key Location/Qualifiers  
 FT Misc-difference 65 /note= "Val in sequence of AAY06566"  
 FT Peptide 566..570  
 FT /note= "conserved element, Asp residues function in  
 coordination of a divalent cation"  
 XX WO9937139-A1.  
 XX 29-JUL-1999.  
 XX 21-JAN-1999; 99WO-US001300.

Thu Aug 26 08:27:51 2004

PR	22-JAN-1998;	98US-0072204P.	
XX	(UNIW )	UNIV WASHINGTON STATE RES FOUND.	
PA	Croteau RB,	Crock JE, Bohlman J, Jetter R, Steele CL;	
PI	WPI; 1999-458569/38.		
XX	N-PSDB; AAX87504.		
DR	Grand fir (Abies grandis) sesquiterpene synthase genes and protein products, useful for sesquiterpenoid production.		
XX	Claim 19; Page 71-74; 136pp; English.		
XX	This sequence represents the E-alpha-bisabolene synthase of grand fir (Abies grandis), a wound-inducible enzyme capable of generating multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl diphosphate. Another deduced sequence (see AAY06566) for the enzyme has Val at position 65. The invention provides gymnosperm sesquiterpene synthase nucleic acids and recombinant proteins, as well as expression vectors, and a method of imparting or enhancing production of a gymnosperm sesquiterpene synthase in a prokaryotic or eukaryotic host cell, especially in a plant, in order to impart, enhance or modify production of sesquiterpenoids, and hence to alter the flavour or fragrance of plant products, to improve disease resistance or to alter ecological interactions mediated by farnesyl diphosphate and its derivatives		
XX	Sequence 817 AA;		
SQ	Query Match 45.5%; Score 2114.5; DB 2; Length 817; Best Local Similarity 49.5%; Pred. No. 3e-182; Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;		
QY	79 RLMADYHPAVKDDFDISLTSNHSATSSVDETINKRIQTLVKEIQCMFQS--MGDGE 136		
DB	25 RRTANPHPNVWGDVHSLKSP-----YIDSSYRERAEVLVSEIKAMLPNPAITGDGE 76		
QY	137 T--NPSAYDTAWARIPSDIGSAPOTFLOILNNQLPDGSGWGBECIFLAYDRVLNTL 194		
DB	77 SMITPSAYDTAWARVPAIDGSRAPQPTQVDMILAKNLKDGSGWGIQSHFLLSDRLATL 136		
QY	195 ACLLTKIKNNKGDIOVKQGVFVRKMEEMKORAD-NHRPSGEVFPVPAMLDEAKSLGID 253		
DB	137 SCVLLKMNVDGLQVEQIEFTKSNLELVKEDTDQDLSVTDFEIIFFPSLLREAAQSRLG 196		
QY	254 LPYHLPISTIOHOKROKLOKIPLVLNHNOTALLYSLGLQDWDWQEIITNQSRDGSF 313		
DB	197 LPYDLPYIHLQTKRQERLAKLSREEIYAVPSLLVSLGIIQDIWEWIMEVQSODGSF 256		
QY	314 LSSPASTACVFMHTQNKRCILHFLNVLKFGDYVPCHPYLDLFLERLWAVDTVERLGDYR 373		
DB	257 LSSPASTACVFMHTGDGAKCLEFLSNVMIKFGNFVPCLPYVDLLERLLIVDNLVRLGIYRH 316		
QY	374 FKXIKESLDYVRYWDAERGVAWCNPIPDVDDTANGRLIRLLHGYNVSSDVLENFRD 433		
DB	317 FEKEIKEALDYVRYHWN-ERGIGRNLNPLADLETTALGFRLLKRLHRYNVSPALFDNFKD 375		
QY	434 EKGDFFCFAGTQIGVTDNLNLYCSQVCPGKIMEEAKFTFTNHLQNALAKNNAFDKW 493		
DB	376 ANGFIKSTGQFNKDVASMLNLYRASQLAFGENILDEAKSFATKYLREALKSETSSAW 435		
QY	494 AVKKDLPGVEVYAIKPYMHRSMPLREARSYTEQGSNDVWLGKTVYKMLVSNKYLELA 553		
DB	436 NNKQNLQSEIKYALKTSWASVPRVEAKRYCQVYRDPYARIKACVYKLPVYNNKFFLELG 495		
QY	554 KLDENWQALHOKETQHLVSNRRSSGENDLFTQRVPMYFVSVAWSFEPEFAACRIAY 613		
DB	496 KLDENIIOSHOEENKWNITSFRUSGLPLDFTFAERPLEFYLVAAGTYEPQYAKCRFLP 555		
QY	614 AKTSCLAVIDDLYDTHGSLDDLKLFSEAVRRWDISVLDSVRDNLQKVCFLGLINTNGF 673		
DB	556 TKVACLQTVLDDMYDTYGTDLDELKLFTEAVRRWDLSTFENLPD-YMKLCYQIYYDIVHEV 614		
QY	674 GKDGKQGRDVLGYLRLKRWEGLLASVYTKAEWSAAKYVTFNFEYVENAKVSIALATVVL 733		
DB	615 AWEAKEQGRVLSVFRKGMEDYLLGYEBAEWLAAYEPTLDYIKNGITTSIGQRILL 674		
QY	734 NSI-FFTGELLPDYILQOVDL--RSKFLHLVSLTGRLLINDTKTYAERNRGELVSSVQCY 790		
DB	675 SGVLMDGQLLSQSALEKVDYPPGRVLTSLNLSRLSLRLADDTKYTKAEKARGELASSIECY 734		
QY	791 MRENPECTEEALSHVYGIIDNALKELNLANPASNAPLCVRLLFNFTARVMQLFYMYR 850		
DB	735 MKDHPCTEEALDHIYSILEPAVKELTRELPK-DDVPFACKQMLFEETRTVTWVIFKDG 793		
QY	851 DFGIGSKEMKDHVSRFLFDVPV 872		
DB	794 DFGVSKLEVNDHIKECLIEPL 815		
XX	RESULT 7		
ID	ABU03781		
XX	ABU09781 standard; protein; 817 AA.		
AC	ABU09781;		
XX	06-AUG-2003 (first entry)		
DE	White fir E-alpha-bisabolene synthase.		
XX	Levopimaradiene synthase; ginkgolide; enzyme; Escherichia coli;		
KW	Candida albicans; Kluyveromyces lactis; Saccharomyces; white fir;		
KW	E-alpha-bisabolene synthase.		
OS	Abies grandis.		
PN	US2002164736-A1.		
XX	07-NOV-2002.		
PF	07-JAN-2002; 2002US-00041007.		
PR	05-JAN-2001; 2001US-0259981P.		
XX	(MATS/) MATSUDA S P T.		
PA	(SCHE/) SCHEPMANN H G.		
XX	Matsuda SPT, Schepmann HG;		
PI	WPI; 2003-352584/33.		
XX	New purified and isolated nucleic acid sequence encoding a levopimaradiene synthase useful for producing pure ginkgolides in cells.		
PT	Example 3; Fig 4; 37pp; English.		
XX	The invention relates to a purified and isolated nucleic acid sequence encoding a levopimaradiene synthase. The nucleic acid is used for the preparation of a ginkgolide in a cell, an expression vector, a unicellular organism (e.g. Saccharomyces, Escherichia coli, Candida albicans or Kluyveromyces lactis), a host cell, a transgenic plant or a seed (particularly ginkgo biloba), a host cell, a transgenic plant and in the molecular biology, for the biosynthesis of ginkgolides and in molecular biology, molecular genetics, and organic chemistry. The nucleic acid is capable of producing large amounts of pure ginkgolides in a cost-effective manner, particularly in an organism capable of a high-yield ginkgolide-producing system. This sequence represents an E-alpha-bisabolene synthase polypeptide which shows similarity to the levopimaradiene synthase of the invention		
XX	Sequence 817 AA;		
SQ	Query Match 45.5%; Score 2114.5; DB 6; Length 817; Best Local Similarity 49.5%; Pred. No. 3e-182; Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;		



QY	494	AVKKDLPEGEVYAIKYPWHRSPRLRSEARSYIEQFGSDNVLGKTVYKMLYVSNKYLELA	553
DB	436	NNQNLSQEIYALAKTSWHASVPRVEAKRYCQVPRPDYARIATAKCYKLPYVNNKFLBGL	495
QY	554	KLDFNMVQALHOKETOHIWVSWRESGFNDLFTFTRQRPVEMVFSVAVSMPEPEFAACRIAY	613
DB	496	KLDFNIIQSIHQEEMKNVTSWFRDSDLPLFTFARERPLEFVPLVAAAGTYEPQYAKCRFLF	555
QY	614	AKTSCLAVIDLDLYDTHGSLDLKLFSEAVRRWDISVLDSVRDNQLKVCFLGLYNTVNGF	673
DB	556	TKVACLQTVLDMDYDYGILDELKLFTEAVRRWDLSTENLPD-YMKLCYQIYYDIVHEV	614
QY	674	GKDLKEQGRDVLGYLRKQWEGLLASYTKEAEWSAAKYVPTFNEVENAKVSIALATVVL	733
DB	615	AWEAEKEQGRDLVSFFRKGMEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGORILL	674
QY	734	NSI-FFTGELLDPYILQQVDL--RSKFLHLVSLTGRLLINDTKTYQAEARNRGLVSSVQCY	790
DB	675	SGVLMDGQLLSQEALEKVDYFGRRVLTENLSLSRLADDTKYKAEKARGELASSIECY	734
QY	791	MRENPECTEAEALSHVYGIIDNALKELNWEANPASNAPLCVRRLLFNTRVMOFLFYMYR	850
DB	735	MKHDPCTEAEALDHIYSILEPAVKELTRBFLKP-DDVPFACKMQLFEETRTVTWVIFKDG	793
QY	851	DGFGISDKEMKHVSRTLDPV 872	
DB	794	DGFGVSKLEVKDHIKECLIEPL 815	
RESULT 9			
AA	AY06568	standard; protein; 817 AA.	
AC	AY06568;		
DT	08-OCT-1999	(first entry)	
DE	E-alpha-bisabolene synthase computer-generated sequence.		
KW	E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;		
KW	gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.		
OS	Synthetic.		
OS	Abies grandis.		
PH	Key	Location/Qualifiers	
FT	Peptide	566..570	
FT	/note= "conserved element, Asp residues function in		
FT	coordination of a divalent cation"		
XX	W09937139-A1.		
XX	29-JUL-1999.		
XX	21-JAN-1999;	99WO-US001300.	
XX	22-JAN-1998;	98US-0072204P.	
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.		
XX	Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;		
XX	WPI; 1999-458569/38.		
XX	N-PSDB; AAX87531.		
XX	Grand fir (Abies grandis) sesquiterpene synthase genes and protein		
PT	products, useful for sesquiterpenoid production.		
XX	Example 11; Page 104-106; 136pp; English.		
CC	This is the sequence of an E-alpha-bisabolene synthase polypeptide		
CC	encoded by a computer-generated nucleic acid (see AAX87531) based on an		

CC	isolated grand fir (Abies grandis) E-alpha-bisabolene synthase clone (see	
CC	AAX87504). It contains conservative amino acid substitutions relative to	
CC	the native sequence (see AY06562). E-alpha-bisabolene synthase is a	
CC	wound-inducible enzyme capable of generating multiple sesquiterpenes,	
CC	especially E-alpha-bisabolene, from farnesyl diphosphate. The invention	
CC	provides gymnosperm sesquiterpene synthase nucleic acids and recombinant	
CC	proteins, as well as expression vectors, and a method of imparting or	
CC	enhancing production of a gymnosperm sesquiterpene synthase in a	
CC	prokaryotic or eukaryotic host cell, especially in a plant, in order to	
CC	impart, enhance or modify production of sesquiterpenoids, and hence to	
CC	alter the flavour or fragrance of plant products, to improve disease	
CC	resistance or to alter ecological interactions mediated by farnesyl	
CC	diphosphate and its derivatives	
XX	Sequence 817 AA;	
QY	Query Match	45.5%; Score 2113.5; DB 2; Length 817;
DB	Best Local Similarity	49.5%; Pred. No. 3.7e-182;
DB	Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;	
QY	79 RLNADYHPAVWKDDFIDSLTSPNSHATSKSVDETINKRIOTLVKEIQCMPOS--MGDGE	136
DB	25 RRTANPHENVNGYDLVHSLKSP-----YIDSSYRERAEVLSEIKAMLNPAITGDGE	76
QY	137 T--NPSAYDTAWARIPSIDSGAPQFPOTLQWILNNOLPDGSGECCIFLAYDRVLNTL	194
DB	77 SMITPSAYDTAWARVPAIDGSRPQFPOTDMLKQNKQKDSWGIQSHFLSLDRLLATL	136
QY	195 ACLLTJKIWNKGDIOVKQGVFVRKHEEMKDEAD-NHRPSGFVYVFPAMLDEAKSLGUD	253
DB	137 SCVIVLLKNVGDQLVQEQGIEFIKTNLELVKDETDDQSLVTDFTFIFFSLLEAQSLRG	196
QY	254 LPYHLPTFSIQHQRKQKQLQKIPLVNLHNHTALYSLLEGLOVDVQWQITNLQSDGSF	313
DB	197 LPYDLPIYHLLQTKRQLAKLSREEIYAVSPPLYSLEGIQDIVEWEHIMEVQSDGSF	256
QY	314 LSSPASTACVFMHTQKRCLEHFLNLFVLSKFGYVPCHPVPLDIFERLWADVTYVERLGIDRY	373
DB	257 LSSPASTACVFMHTGDGAKLEFLNSVMKFGNFVCLYFPVLLERLLIVDNIVRLGIYRH	316
QY	374 FKKEIKESLDYVYVYWDAGRGVGMARCNPIPDVDTAMGLRLILRLHGYVNSDVLNFRD	433
DB	317 FEKEIKEALDYVYVYVHWN-ERGIWGRNLNPIADLETTALGFRLLRLHRYNVPAINFKD	375
QY	434 EKGDFCFPAGOTQIGYTDNLNLYRCQVCPGKIMEEAKTFTTNHLONALAKNNAFDKW	493
DB	376 ANGKFCSTGQFNKDVASMLNLYRASQLAPGENILDEAKSFATKYLREALEKSETSSAM	435
QY	494 AVKKDLPEGEVYAIKYPWHRSPRLRSEARSYIEQFGSDNVLGKTVYKMLYVSNKYLELA	553
DB	436 NNQNLSQEIYALAKTSWHASVPRVEAKRYCQVPRPDYARIATAKCYKLPYVNNKFLBGL	495
QY	554 KLDPMVQALHOKETOHIWVSWRESGFNDLFTFTRQRPVEMVFSVAVSMPEPEFAACRIAY	613
DB	496 KLDPMVQALHOKETOHIWVSWRESGFNDLFTFTRQRPVEMVFSVAVSMPEPEFAACRIAY	555
QY	614 AKTSCLAVIDLDLYDTHGSLDLKLFSEAVRRWDISVLDSVRDNQLKVCFLGLYNTVNGF	673
DB	556 TKVACLQTVLDMDYDYGILDELKLFTEAVRRWDLSTENLPD-YMKLCYQIYYDIVHEV	614
QY	674 GKDLKEQGRDVLGYLRKQWEGLLASYTKEAEWSAAKYVPTFNEVENAKVSIALATVVL	733
DB	615 AWEAEKEQGRDLVSFFRKGMEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGORILL	674
QY	734 NSI-FFTGELLDPYILQQVDL--RSKFLHLVSLTGRLLINDTKTYQAEARNRGLVSSVQCY	790
DB	675 SGVLMDGQLLSQEALEKVDYFGRRVLTENLSLSRLADDTKYKAEKARGELASSIECY	734
QY	791 MRENPECTEAEALSHVYGIIDNALKELNWEANPASNAPLCVRRLLFNTRVMOFLFYMYR	850
DB	735 MKDHPECTEAEALDHIYSILEPAVKELTRBFLKP-DDVPFACKMQLFEETRTVTWVIFKDG	793
QY	851 DGFGISDKEMKHVSRTLDPV 872	





XX WPI; 1999-120396/10.  
DR N-PSDB; AAX08654.  
XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand fir (Abies grandis), used to provide plants with modified production of  
PT monoterpenes, e.g. myrcene, limonene or pinene.  
PT  
XX  
XX Example 1; Page 90-92; 12lpp; English.  
XX Nucleotide sequences encoding myrcene synthase, limonene synthase and  
CC pinene synthase from Grand fir may be incorporated into any organism  
CC (e.g. intact plant, animal, microbe), or derived cell culture that  
CC produces geranyl diphosphate for the production of the aforementioned  
CC enzymes or their products. The sequences when expressed in transfected  
CC cells may also be used for the production or modification of flavour and  
CC aroma properties, improvement of defense capability, and the alteration  
CC of other ecological interactions mediated by myrcene, limonene, pinene,  
CC or their derivatives. In particular they can be used for the production  
CC of plant seeds for the extraction of oil or meal. cDNAs encoding myrcene  
CC synthase, pinene synthase and limonene synthase from Grand Fir were  
CC isolated and sequenced by the following procedure. Based on comparison of  
CC sequences of different plant synthase enzymes, four conserved sequences  
CC were identified for construction of four degenerate primers (AAX08646-  
CC 49). One of these primers (AAX08648) was an antisense primer and used for  
CC PCR with the other three sense primers. Only a combination of primers  
CC AAX08648 and AAX08649 generated a PCR product which was 110 base pairs in  
CC length. The amplified product was used in a plasmid construct to  
CC transform E.coli XL1-Blue cells and the inserts were sequenced. Four  
CC probes (AAX08650-53) were devised from the sequenced insert and were used  
CC to screen a Grand Fir cDNA library. One of these clones is the myrcene  
CC synthase sequence given in AAX08643. The other clones identified are  
CC described in AAX08654-56  
XX  
SQ Sequence 782 AA;  
Query Match 44.9%; Score 2085.5; DB 2; Length 782;  
Best Local Similarity 49.6%; Pred. No. 1.2e-179;  
Matches 391; Conservative 158; Mismatches 221; Indels 19; Gaps 9;  
52 DFTDLSLTPNSHATSKSSVDTEINKRIQTIVKIEIQCFQS--MGDGET--NPSAYDTAW 147  
3 DLVHLSKSP-----YIDSSYREAEVLVSIIKVMNLPAITGDSMTTPAYDTAW 54  
148 ARIPSIDGSGAPPPOTLWLNQLPDGSGBEICIFLAYDRVLNTLACLLTKIKNKG 207  
55 ARVPAIDGSRAPQPPQTVDMILNQLKDGSGIQQSHFLLSDRLATLSCVLVLLKNVGD 114  
208 IQVQGVFVRKGMMEEMKDEAD-NHRPSGEVFPVPPAMLDKASIGLDLPYHLPFIQIHQ 266  
115 LQVEQGEIFKSNLELVKDETDQDSLVTDPEIIPPSLLRQAQSLRGLPYDLPYHLLQT 174  
267 KRQKKLOKIPLVLNHNTALLYSLGLQVDWDOEITNLQSRDGLSPSSPASTACVFMH 326  
175 KRQRLAKLGRREIYAVPSPLLYSLEGIQDIVERIMVQSDGSLSPASTACVFMH 234  
327 TQNKRCILHFLNFKSGDYVPCHPYLDLPERLWAVDTVERLGDIDRYFKKEIKESLDYV 386  
235 TGDACKLEFLNSVMKFGNFPVCLPYDILLERLIVDNLVRLGIYRHFKEIKESLDYV 294  
387 RYWDARGVGWACNPIDVDVDTANGLRILRLHGVNYSVDLENFRDEKGFDFCFAGQTQ 446  
295 RHWN-ERGIWGRLNPLNLTALGFLRLKLRHNVSPALFDFKQANGKFCSTQGFN 353  
447 IGVTDLNLVRCQVCPGGEIMEEAKFTTNHNLQNALAKNAPDKWAVKQDLPGVEYEA 506  
354 KDVASMLNLYRASQLAPFGENILDEAKSFATKYLREALKESSETSSAWNKNQSLQIKYA 413  
507 IKYPHWRSMPLERARSYIEQSGNDVWGKTVYKMLYVSNKYLELAKLDLMMVQALHOK 566  
414 LKTSWHASVPRVEAKRYQVVRPDYAKVCYKLPYVNNKFLGKLDLNNIQLSHQ 473  
567 ETQHIVSWWRSGFNDLFTTRQRPVMEYFVAVSNMFEPEFAACRIAYAKTSCLAVILDDL 626

Db 474 EMKNVTISWFRDGLPFLTFEAREPFLFYFLVAAAGTYEPQYAKCRFLFTKVAQLQVLDDM 533  
Qy 627 YDTHGSLDDDLKLFSEAVRRWDISVLDSVRDNLKVCFLGLYNTVNGFGKGLKEQGRDVL 686  
Db 534 YDTGTDELKLFTEAVRRWDLSTENLPD-YMKLCYQIYYDIVHEVAWEAEKEQGRBLV 592  
Qy 687 GYLKRWEGLLASYTKAEWSAAKYVPTFNFNENAKVSIATATVVLNSI-FFTCELLPD 745  
Db 593 SFFRKGWEDYLLGYVEEAELAAEYVPTLDEYIKNGITSIGORILLSGVLMDGQLLSQ 652  
Qy 746 YILQOVDL--RSKFLHLVSLTGRLLNDKTYOASNRGELVSSVOCYMRNPECTEERAL 803  
Db 653 EALEKVDYPSGRVLTETLSLSRLADDTKTYAEKARGELASSIECYMKDHPCTEERAL 712  
Qy 804 SHVGIIDNALKEINLWELANPASNAPLCVRRLLENTARVMQLFYMYRDPGFGISDKEMKDH 863  
Db 713 DRIYSILSPAVKELTREFLKP-DDVPPACKMKWLPFEETRTVMVFKDGGDFGVSKLEVKDH 771  
Qy 864 VSRTLFDPV 872  
Db 772 IKECLIEPL 780  
RESULT 12  
AAX08653  
ID AAX08653 standard; protein; 782 AA.  
XX AAX08653;  
XX  
XX 25-AUG-2000 (first entry)  
XX Grand fir E-alpha-bisabolene synthase protein sequence SEQ ID NO:46.  
XX  
XX Synthase; protein co-ordinate data; active site; modification; terpenoid;  
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
KW defensive agent; pigment; antitumour; steroid hormone;  
KW signal transduction pathway; bile acid; affinity purification;  
KW photoreceptor; enzymatic synthesis; nutrient supplement;  
KW immunological reagent.  
XX  
XX Abies grandis.  
XX  
XX WO200017327-A2.  
XX  
XX 30-MAR-2000.  
XX  
XX 17-SEP-1999; 99WO-US021419.  
XX  
XX 18-SEP-1998; 98US-0100993P.  
XX 22-APR-1999; 99US-0130628P.  
XX 23-AUG-1999; 99US-0150262P.  
XX  
XX (KENT ) UNIV KENTUCKY RES DEPT.  
XX (SALK ) SALK INST BIOLOGICAL STUDIES.  
XX  
XX Chappell J, Manna KR, Noel JP, Starks CM;  
XX  
XX WPI; 2000-292839/25.  
XX N-PSDB; AAA38932.  
XX  
XX Novel terpene synthase enzymes, useful for producing terpene  
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
PT enzymes by specific amino acid alterations.  
XX  
XX Claim 108; Page 418-420; 450pp; English.  
XX  
XX The present invention describes an isolated terpene synthase (I)  
CC comprising a region with at least 20% identity to region 265-535 of a 548  
CC amino acid (aa) sequence (Ia), given in AAX08631. (I) contains nine alpha  
CC -carbon atoms (alphaC) that have interatomic distances, between each  
CC other, within tabulated ranges, have a centre point (within a sphere of



```

Db 175 KRQERLAKLSREIYAVPSPLLYSLGDIQIVEMERIMEVQSDGSLSSPASTACVFMH 234
QY 327 TQNKRCCLHFLNVLKFGDYVPCYHPLDLPERLMAVDTVRLGIDRYFKFKIKESLDYVY 386
Db 235 TGDACKLEFLNSVMIKFGNFVPCLYPDVLLERLLIVDNIIVRLGIYRHFKEIKEALDYVY 294
QY 387 RYWDAERGVMGWCNPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEKDFCFAGOTQ 446
Db 295 RHWN-ERGIWGRNLNPAIDLETTALGFRLLRLHRYNVSPAIFDNFKDANGKFCSTGQFN 353
QY 447 IGYTDNLNLYRCSQVCPGGEKIMEEAKTFTTNHLQNALAKNNAFADKWAVKDLPGEVEYA 506
Db 354 KDVASMLNLYRASQLAFPGENILDEAKSPATKYLREALEKSETSSAANNKQLSQEIKYA 413
QY 507 IKYPHRSMPRLRARSYIBQFGSDNVWLKGTVTYKMLYVSNKYLELAKLDFNMVQALHOK 566
Db 414 LKTSWHSASVPRVEAKRYCQVYRDPYARIACVYKLPYVNNKEFLELGLDFNIIQSIHQE 473
QY 567 ETQHIYSWWRRESGFNDLTFTQRPVMEYFSAVSMPEPEFAACRIAYAKTSCLAVLDDL 626
Db 474 EMKNVTSWFRDGLPLFTFARERPLEFYFLVAAAGTYEPQYAKRFLTKVACLOTVLDDM 533
QY 627 YDTHGSLDDLKLFSEAVRRWDISVLDSVRDNLKVCFLGLYNTVNGFGKDKLKEQGRDVL 686
Db 534 YDTYGTLDLKLFTAEAVRRWDLSTENLPD-YMKLCYQIYVDIVHEVAWEAKEQGRLEV 592
QY 687 GYLKRWEGLLASYTKEAWEASAAYPTFTFNEYVENAKVSIATATVVLNSI-FFTGELLDP 745
Db 593 SPFRKGWEDYLLGYBEAEWLAAYEYPTLDEYIKNGITSIGQRILLSGVLIMDQGLLSQ 652
QY 746 YILQOVDL--RSKFLHLVSLTGRINDTKTYQABERNRDELVSVCYCMRENPECTEELAL 803
Db 653 EALEKVDYPCRRVLTENSLISRLADTKTYKAEKARGELASSIECYMKDHPCTEELAL 712
QY 804 SHYGIIDNALKEINWELANPASNAFLCVRRLLNTARVNLQPLFYMYRDFGFGISKEMKOH 863
Db 713 DHIYSILEPAVKELTEFLKP-DDVFPACKMLFEETRTVTWIFKDGDFGVSKLEVKDH 771
QY 864 VSRITLDPV 872
Db 772 IKECLIEPL 780

RESULT 14
ABB79392
ID ABB79392 standard; protein; 782 AA.
XX
AC ABB79392;
XX
DT 13-AUG-2002 (first entry)
XX
DE Taxadiene synthase homologous protein sequence w85703.
XX
KW Taxane; taxane synthase; taxadiene synthase; cytostatic; anticancer;
KW enzyme.
XX
OS Unidentified.
XX
PN WO200240694-A2.
XX
PD 23-MAY-2002.
XX
PF 16-NOV-2001; 2001WO-DK000763.
XX
PR 17-NOV-2000; 2000DK-00001730.
XX
PR 23-NOV-2000; 2000US-0253843P.
XX
PA (NOVO ) NOVOZYMES AS.
XX
PI Vind J;
XX
XX WFI; 2002-471624/50.
XX

```

```

PT Heterologous production of taxane analogs related compound useful as
PT anticancer agent involves cloning a full-length taxane synthesis pathway
PT from a taxan-producing organism into a taxane-resistant host cell.
XX
XX Example 3; Fig 1; 60pp; English.
XX
CC The present invention describes the heterologous expression of taxane
CC analogues, comprising cloning a DNA sequence comprising a taxane
CC synthesis pathway, making a DNA construct in which the DNA sequence is
CC under control of regulatory elements, introducing the DNA construct into
CC a host cell, growing the host cell to produce taxane analogues, and
CC recovering the required taxane from the culture medium. Also described:
CC (1) an isolated DNA sequence comprising the taxane synthesis pathway,
CC (preferably taxol synthesis pathway); (2) an expression vector comprising
CC the isolated DNA sequence; and (3) a host cell comprising a taxane
CC synthesis pathway derived from a taxane-producing organism, plant or
CC tree, the taxane synthesis pathway is foreign to the host cell. The
CC taxane analogues have cytostatic activity, and so can be used as
CC anticancer agents. The present sequence represents a taxadiene synthase
CC homologous protein, which is given in an example from the present
CC invention
XX
SQ Sequence 782 AA;
Query Match 44.9%; Score 2085.5; DB 5; Length 782;
Best Local Similarity 49.6%; Pred. No. 1.2e-179;
Matches 391; Conservative 158; Mismatches 221; Indels 19; Gaps 9;
QY 92 DRIISLTSNHSATSKSSYDTERINKRIQTLVKEIQCMFOS--MGDGET--NPSAYDTAW 147
Db 3 DLVHSLKSP-----YIDSSYREAEVLVSEIKVMLNPAITGDSGEMITPSAYDTAW 54
QY 148 ARIPSDIGGAGPOFOTLQWILNNQLPDGSGWBECEIFLAYDRVLNLTACLTLKWNKGD 207
Db 55 ARVPAIDGARGQFPFTVDWILKNQLKDGSGWGLQSHFLLSDRLATLSCVLLKWNVGD 114
QY 208 IQVQKGVFVRKHEEMKDEAD-NHRPSGFVVFPAMLEAKSLGLDLPLHLFPISQIHQ 266
Db 115 LQVEQIEFIKSNLELVKDETDQDSLVTDTPTFIIFPSLLREAOQLRLGLPYLPIHLQ 174
QY 267 KRQKLOKIPLVNLNHNQTALLYSLGLQDVVDQWQITNLQSRDGSFLSSPASTACVFMH 326
Db 175 KRQERLAKLSREIYAVPSPLLYSLGDIQIVEMERIMEVQSDGSLSSPASTACVFMH 234
QY 327 TQNKRCCLHFLNVLKFGDYVPCYHPLDLPERLMAVDTVRLGIDRYFKFKIKESLDYVY 386
Db 235 TGDACKLEFLNSVMIKFGNFVPCLYPDVLLERLLIVDNIIVRLGIYRHFKEIKEALDYVY 294
QY 387 RYWDAERGVMGWCNPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEKDFCFAGOTQ 446
Db 295 RHWN-ERGIWGRNLNPAIDLETTALGFRLLRLHRYNVSPAIFDNFKDANGKFCSTGQFN 353
QY 447 IGYTDNLNLYRCSQVCPGGEKIMEEAKTFTTNHLQNALAKNNAFADKWAVKDLPGEVEYA 506
Db 354 KDVASMLNLYRASQLAFPGENILDEAKSPATKYLREALEKSETSSAANNKQLSQEIKYA 413
QY 507 IKYPHRSMPRLRARSYIBQFGSDNVWLKGTVTYKMLYVSNKYLELAKLDFNMVQALHOK 566
Db 414 LKTSWHSASVPRVEAKRYCQVYRDPYARIACVYKLPYVNNKEFLELGLDFNIIQSIHQE 473
QY 567 ETQHIYSWWRRESGFNDLTFTQRPVMEYFSAVSMPEPEFAACRIAYAKTSCLAVLDDL 626
Db 474 EMKNVTSWFRDGLPLFTFARERPLEFYFLVAAAGTYEPQYAKRFLTKVACLOTVLDDM 533
QY 627 YDTHGSLDDLKLFSEAVRRWDISVLDSVRDNLKVCFLGLYNTVNGFGKDKLKEQGRDVL 686
Db 534 YDTYGTLDLKLFTAEAVRRWDLSTENLPD-YMKLCYQIYVDIVHEVAWEAKEQGRLEV 592
QY 687 GYLKRWEGLLASYTKEAWEASAAYPTFTFNEYVENAKVSIATATVVLNSI-FFTGELLDP 745
Db 593 SPFRKGWEDYLLGYBEAEWLAAYEYPTLDEYIKNGITSIGQRILLSGVLIMDQGLLSQ 652
QY 746 YILQOVDL--RSKFLHLVSLTGRINDTKTYQABERNRDELVSVCYCMRENPECTEELAL 803

```

```

Db 653 EALEKVDPGRRVLTENSLISRLADDTKTKAKARGELASSTECTMKDHPCTEEAL 712
Qy 804 SHVYGIIDNALKELNWLNPASNAPLCVRRLLFNTRVNMQLFMYRDFGIGSDKEMKH 863
Db 713 DHIYSILEPAVKELTREFLKP-DDVPPACKMKLFEETRTVTMTVIFKDGFGVSKLEVKH 771
Qy 864 VSRTLFDPV 872
Db 772 IKSECLIEPL 780

RESULT 15
AAW31655
ID AAW31655 standard; protein; 862 AA.
AC AAW31655;
DT 30-APR-1998 (first entry)
XX Pacific yew taxadiene synthase.
DE Pacific yew; taxadiene synthase; taxol biosynthesis; probe; immunoassay.
KW Taxus sp.
OS
XX WO9738571-A1.
XX 23-OCT-1997.
XX 15-APR-1997; 97WO-US006320.
XX 15-APR-1996; 96US-0015993P.
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX Wildung MR, Croteau RB;
XX WPI; 1997-526123/48.
XX N-PSDB; AAT97447.
XX DNA encoding Pacific yew taxadiene synthase - used to increase taxol biosynthesis.
XX Disclosure; Fig 2; 30pp; English.
XX This is a cloned pacific yew taxadiene synthase (TS) protein. The sequence can be used to increase to increase taxol biosynthesis into taxoids in transformed cells. The protein can be modified to improve efficiency, and to redirect the targeting of the expressed polypeptide. The modifications can be to alter catalytic efficiency, to modify substrate utilisation, and to alter product outcome. The antibodies specific for the mature native pacific yew TS polypeptide can be used to purify TS polypeptides, cloning TS homologues from Pacific yew or other plant species, and as probes for blots and immunoassays

SQ Sequence 862 AA;
Query Match 42.2%; Score 1959; DB 2; Length 862;
Best Local Similarity 43.6%; Pred. No. 4.3e-168;
Matches 385; Conservative 167; Mismatches 255; Indels 76; Gaps 15;

Qy 33 KRSSFQNAQH-----C-VRSHLRLRWC-----VGIIHAS 61
Db 11 KNNALGNKAHDPNCRACKSERQMMVCSRSGRTRVKMSRGSGGPGPVMMSSSTGTSKV 70
Qy 62 AAETRP---DOLPOEERFVRLNADYHPAVWKDDFIDSLTSPNSHATSKGSVDETINKRI 118
Db 71 VSETSTIVDDIP-----RLSANYHGDLHHRNVIQTLTPFPRESS-----TYQERA 116
Qy 119 QTLVKEIQCFQSGMGDGTNPASVDTAWARIPSI--DGS GAPQFPQTLOWILNNQLPDG 176
Db 117 DELVVKIKDMFNALGDGDIDSPSADYTAWARLATISSDSEKPRFPQALNWFNNQLQDG 176

```

```

Qy 177 SWGECIFLAYDRVLTNTLACLTLKIWNKGDIQVKGVEFVRKHEEMKDEADNHRPSGF 236
Db 177 SWGIEHFSLCDRLNTNTNSIALSVMTKGTSHSQVGAEFIAENLRLLNEE-DELSPD-F 234
Qy 237 EVVFPAMLDEAKSIGLDLPVHLPPFISQIHOKROKLL-----QKIPLNVLHNHOTALLY 289
Db 235 QIIFPALLQKAKALGINLPYDLFFIKYLSTTRARLTDVSAADNIPANMLN----- 286
Qy 290 SLEGLQVVDWQBITNLQSRDGSFLSSPASTACVFMHTONKRCILHFLNVLKFGDVPC 349
Db 287 ALEGLEVIDWNKIMRFQSKDGSFLSSPASTACVLMNTGDEKCFTLNLLDKFGGCVPC 346
Qy 350 HYPDLDFERLWADVTVERLIGIDRYFKKEIKESLDYVYRWDAERGVCWACNPDPVDDT 409
Db 347 MYSIDLRLSLVDNTEHLGIGRHFQKEIKGALDYVYRHW-SERGIQGWGSDSLVPDLNTT 405
Qy 410 AMGLRILRLHGVNVSVDLENFRDEKDPFCFAGTOIGVTDNLNLNLCVRCQVCPGSKIM 469
Db 406 ALGLRLRMHGVNVSVDLENFRDENGFRPFSSAGQTHVELRSVVNLFRASDLAPPDERAM 465
Qy 470 EEAKTFTTHLQNALAKNNAFQKWA VKCOLPGEVEYAIKYPMHRSMRPLEARSYIEQFGS 529
Db 466 DDARKFAEPYREALA-----TKISTNTKLFKEIYVVEYPMHMSIPRLEARSYIDSYDD 520
Qy 530 NDVWLKGTVMKLYVSNKYLELAKLDFNNVQALHOKETOHIIVSWWRESGFNDLTFTQR 589
Db 521 NYVWQRKTLYRMPSLNSKCLELAKLDFNTVQSLHQBELKLLTRWKKESGMADINFTRHR 580
Qy 590 PVEMVFSVAVSMFEPFPAACRIAYAKTSLAVILDLDLYDTHGSLDDLKLFSEAVRRWDIS 649
Db 581 VAEYFSSAT--PEPEYSATRIAFTKIGLQVLPDDMADIFATLDELKSFTEGVKRWDT 638
Qy 650 VLDSVRDNLQKVCFLGLYNTVNGFGKDGKLBQGRDVLGYLRKVMWEGLLASYTKEAESA 709
Db 639 LLHEIPE-CMQTCFVWFKLMEEVNNDVVKVQGRDMLAHIRKPWELYFNCVQBERWLEA 697
Qy 710 KYVPTFNEYVENAKVSIATVVLNSIFFFTGELLPDYILOQVDLRSKFLHLVSLTGLIN 769
Db 698 GYIPTFEYELKTYAISVGLGCPCTLOPILLMGLVKDDVVEKHVYPSNMFLVLSLWLTN 757
Qy 770 DTKTYQAEERNRGELVSSVOCYMRNPECTEEALSHVYGIIDNALKELNWLNPASNA 829
Db 758 DTKTYQAEKARGQQAGIACVMDKNPGATEEDAIKHICRVVDRLAKREASFEYFKPSNDIP 817
Qy 830 LCVRLLFNTRVNMQLFYMYRDFGIGSDKEMKHDSRTLFDPV 872
Db 818 MGCKSFIFNLRLCVQIFKYFIDGYIANEBIKDYIRKVYIDPI 860

```

Search completed: August 17, 2004, 21:00:10  
Job time : 57.4614 secs

**THIS PAGE IS BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:58:21 ; Search time 18.8728 Seconds  
(without alignments)  
2388.070 Million cell updates/sec

Title: US-10-041-007-2

Perfect score: 4645

Sequence: 1 MAGVLFANLPCSLQLSPKVP.....GISKEMKHVSRITLDPVA 873

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgm2\_6/ptodata/2/iaa/5A\_COMB.pap:\*
- 2: /cgm2\_6/ptodata/2/iaa/5B\_COMB.pap:\*
- 3: /cgm2\_6/ptodata/2/iaa/6A\_COMB.pap:\*
- 4: /cgm2\_6/ptodata/2/iaa/6B\_COMB.pap:\*
- 5: /cgm2\_6/ptodata/2/iaa/PCTUS\_COMB.pap:\*
- 6: /cgm2\_6/ptodata/2/iaa/backfiles.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2844	61.2	868	4	US-09-398-395A-56 Sequence 56, Appli
2	2844	61.2	868	4	US-09-887-586A-56 Sequence 56, Appli
3	2844	61.2	868	4	US-09-895-752-56 Sequence 56, Appli
4	2844	61.2	868	4	US-09-903-012B-56 Sequence 56, Appli
5	2844	61.2	868	4	US-09-900-797-56 Sequence 56, Appli
6	2114.5	45.5	817	3	US-09-234-393-13 Sequence 13, Appli
7	2114.5	45.5	817	3	US-09-234-393-40 Sequence 40, Appli
8	2114.5	45.5	817	4	US-09-865-171-13 Sequence 13, Appli
9	2114.5	45.5	817	4	US-09-865-171-40 Sequence 40, Appli
10	2113.5	45.5	817	3	US-09-234-393-38 Sequence 38, Appli
11	2113.5	45.5	817	3	US-09-234-393-42 Sequence 42, Appli
12	2113.5	45.5	817	4	US-09-865-171-38 Sequence 38, Appli
13	2113.5	45.5	817	4	US-09-865-171-42 Sequence 42, Appli
14	2085.5	44.9	782	4	US-09-360-545-16 Sequence 16, Appli
15	2085.5	44.9	782	4	US-09-398-395A-46 Sequence 46, Appli
16	2085.5	44.9	782	4	US-09-887-586A-46 Sequence 46, Appli
17	2085.5	44.9	782	4	US-09-895-752-46 Sequence 46, Appli
18	2085.5	44.9	782	4	US-09-903-012B-46 Sequence 46, Appli
19	2085.5	44.9	782	4	US-09-900-797-46 Sequence 46, Appli
20	1959	42.2	862	3	US-09-315-861-2 Sequence 2, Appli
21	1959	42.2	862	4	US-09-398-395A-44 Sequence 44, Appli
22	1959	42.2	862	4	US-09-887-586A-44 Sequence 44, Appli
23	1959	42.2	862	4	US-09-895-752-44 Sequence 44, Appli
24	1959	42.2	862	4	US-09-903-012B-44 Sequence 44, Appli
25	1959	42.2	862	4	US-09-593-253-2 Sequence 2, Appli
26	1959	42.2	862	4	US-09-900-797-44 Sequence 44, Appli
27	1404	30.2	811	4	US-09-614-912-22 Sequence 22, Appli

28	1332	28.7	802	1	US-08-261-465-2 Sequence 2, Appli
29	1332	28.7	802	1	US-08-405-254-6 Sequence 6, Appli
30	1332	28.7	802	2	US-08-326-286-4 Sequence 4, Appli
31	1318	28.4	823	1	US-08-261-465-1 Sequence 1, Appli
32	1318	28.4	823	1	US-08-405-254-5 Sequence 5, Appli
33	1255	27.0	581	3	US-09-234-393-44 Sequence 44, Appli
34	1255	27.0	581	3	US-09-865-171-44 Sequence 44, Appli
35	1252	27.0	580	3	US-09-234-393-48 Sequence 48, Appli
36	1252	27.0	580	3	US-09-865-171-48 Sequence 48, Appli
37	1252	27.0	581	3	US-09-234-393-20 Sequence 20, Appli
38	1252	27.0	581	3	US-09-234-393-46 Sequence 46, Appli
39	1252	27.0	581	4	US-09-865-171-20 Sequence 20, Appli
40	1252	27.0	581	4	US-09-865-171-46 Sequence 46, Appli
41	1252	27.0	581	4	US-09-398-395A-48 Sequence 48, Appli
42	1252	27.0	581	4	US-09-887-586A-48 Sequence 48, Appli
43	1252	27.0	581	4	US-09-895-752-48 Sequence 48, Appli
44	1252	27.0	581	4	US-09-903-012B-48 Sequence 48, Appli
45	1252	27.0	581	4	US-09-900-797-48 Sequence 48, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-398-395A-56

; Sequence 56, Application US/09398395A

; Patent No. 6468772

; GENERAL INFORMATION:

; APPLICANT: Chappell, Joseph

; APPLICANT: No. 64687721, Joseph P.

; APPLICANT: Starks, Courtney M.

; APPLICANT: Manna, Kathleen R.

; TITLE OF INVENTION: SYNTHASES

; FILE REFERENCE: 07678-025001

; CURRENT APPLICATION NUMBER: US/09/398.395A

; CURRENT FILING DATE: 1999-09-17

; PRIOR APPLICATION NUMBER: 60/100,993

; PRIOR FILING DATE: 1998-09-18

; PRIOR APPLICATION NUMBER: 60/130,628

; PRIOR FILING DATE: 1999-04-22

; PRIOR APPLICATION NUMBER: 60/150,262

; PRIOR FILING DATE: 1999-08-23

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 56

; LENGTH: 868

; TYPE: PRT

; ORGANISM: Abies grandis

US-09-398-395A-56

Query Match 61.2%; Score 2844; DB 4; Length 868;

Best Local Similarity 62.3%; Pred. No. 2.7e-279;

Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;

Qy	12	SLQLSPKVP---FRQSTNLLIPFKRSSFGNACVSRSHLRWN-----CVGIHA 60
Db	5	SSLSUQIPTAAHHLTANAQSIPIHFTTLNAGSSAKRSRSLYLRWGKSNKIIACVG--E 62
Qy	61	SAATRDPQLPQERFVSR---LNADYHPAVWKDDFIDSLTSPNSHATSKSSVDETINKR 117
Db	63	GGATSVYQSAEKNDLSLSSTLVKREPPPGWKDDLDLSLT--SHKVAAS--DB--KR 115
Qy	118	IQTULVKIQCWFQSMGGETNPSAYDTAWARIPIGSGAPQPPQTQLWINNQLPDGS 177
Db	116	IETLISEIKMFCWGYGETNPSAYDTAWARIPIGSGSDNPHFPPTVETWLTQNLKDG 175
Qy	178	WGECIFLAYDRVLTACLLTLKWNKGDIOVKGVFVRKHEEKDEADNRHPSGFE 237
Db	176	WGEFYFLAYDRILATLACITLTLWTGETQVQKGEFFFTQAGKMEDEADSRHPSGFE 235
Qy	238	VVFPAMLDKASGLDLPYHLPPFTSQIHOKRQKQKQKIPNLVNLHNOHTALLYSLEGLQDV 297
Db	236	IVFPAMLEAKILGLDLPYDLFPLKQLEKAKRIPTDVLVALPTLLYSLEGLQEI 295

QY	298	VDWQBITNLQSRDGSFLSSPASTACVFMHTQNKRCFLHFNVLKFGDYVPCHPYPLDLFE	357
Db	296	VDWQIKMKLQSKDGSFLSSPASTAAVFMRTGNKKCLDFLNFVLKFGNHPCHPYPLDLFE	355
QY	358	RLWAVDVTVERLIGIDRYFKKEIKESLDYVRYWDAERGVMARCNDIPDVTDTAMGLRLR	417
Db	356	RLWAVDVTVERLIGIDRHFKKEIKEALDYVYSHWD-ERGIWARENPVDIDDTAMGLRLR	414
QY	418	LHGYNVSSDLNFRDEKDPFCFAGQTOIGVTDMLNLYRCSQVCFPGKEIMEEAKTFTT	477
Db	415	LHGYNVSSDLVLTFRDENGECFCFLGQGTQVTDMLNLYRCSHVSPGGETIMEEAKLCTE	474
QY	478	NHLQNALAKNNADFKWAVKCOLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSNDYWLKGT	537
Db	475	RYLRLNALENVDADFKNWPKKNIRGEVEYALKYPMHKSMPRLSEARSYIENYGPDDVWLKGT	534
QY	538	VYKMLYVSNKYLELAKLDFNMVQALHOKETOHIHVSWMRESGFNDLTFTRQRPVEMYFSV	597
Db	535	VYMPFYISNEKYLELAKLDFNKVQSHQTELOLRWWSKSSGFTDLNFRERVTEIYFSP	594
QY	598	AVSMPEPEPAACRIAYAKTSCILAVILDDLYDTHGSLDDKLFSSEAVRRWDISVLDSDVRDN	657
Db	595	ASPIFEPEPKCREVYTKTSNFTVILDDLYDAHGSLLDKLFTESVKRWDLSDVDQM-PQ	653
QY	658	QUKVCFPLGINTVNGFGKDLKEQGRDVLGYLRKWEGGLASVYTKAEWSAAKYVPTENE	717
Db	654	QMKICFVGFTNFNDIAKEGRERQGRDVLGYIQNVKWKVQLEAYTKAEWSEAKYVPSFNE	713
QY	718	YVENAKVSTALATVNLNSIFFTGELLPDYILOQVDLRSKFLHLVSLTGRINDTKTYQAE	777
Db	714	YIENASVSIALGTVNLISALFTGEVLTDEVLSKIDRESFLQMLGLTGRLVNDTKTYQAE	773
QY	778	RNRGELVSSVQCYMRENPECTEBSALSHVYGIIDNALKEINWELANPASNAPLCVRRLLF	837
Db	774	RQGEVASAIQCYMKDHPKISEEALQHYVSMENALEELNREFVN--NKIPDIYKRLVLF	831
QY	838	NTARVMQLFMYRDGFGIS-DKEMKDHVSRTLPDPVA	873
Db	832	ETARINQLFMYQDGLTSLSHDMEIKHVKVKNCLFPQVA	868
RESULT 3			
US-09-895-752-56			
; Sequence 56, Application US/09895752			
; Patent No. 6559297			
; GENERAL INFORMATION:			
; APPLICANT: Chappell, Joseph			
; APPLICANT: No. 65592971, Joseph P.			
; APPLICANT: Starks, Courtney M.			
; APPLICANT: Manna, Kathleen R.			
; TITLE OF INVENTION: SYNTHASES			
; FILE REFERENCE: 07678-025001			
; CURRENT APPLICATION NUMBER: US/09/887,586A			
; CURRENT FILING DATE: 2001-06-22			
; PRIOR APPLICATION NUMBER: 09/398,395			
; PRIOR FILING DATE: 1999-09-17			
; PRIOR APPLICATION NUMBER: 60/130,628			
; PRIOR FILING DATE: 1999-04-22			
; PRIOR APPLICATION NUMBER: 60/150,262			
; PRIOR FILING DATE: 1999-08-23			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 56			
; LENGTH: 868			
; TYPE: PRT			
; ORGANISM: Abies grandis			
US-09-887-586A-56			
Query Match 61.2%; Score 2844; DB 4; Length 868;			
Best Local Similarity 62.3%; Pred. No. 2.7e-279;			
Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;			



Qy	778	RNRGELVSSVQCYMRENPECTRESEALSHVYGIIDNALKEINWELANPASNAPLCVRRLLF	833
Db	774	RGQGEVASAIQCVMKDHPKISEEALQHVYSYMERALBELNREFVN--NKIPDIYKRLVF	833
Qy	838	NTARVMQLFVYMRDGFIS-DKEMKDHVSRTLFDPA 873	
Db	832	ETARIMQLFVMOGDGLTLSDHMEIKHVKNCILFQPA 868	
RESULT 4			
US-09-903-012B-56			
; Sequence 56, Application US/09903012B			
; Patent No. 6569656			
; GENERAL INFORMATION:			
; APPLICANT: Chappell, Joseph			
; APPLICANT: No. 65696561, Joseph P.			
; APPLICANT: Starks, Courtney M.			
; APPLICANT: Manna, Kathleen R.			
; TITLE OF INVENTION: SYNTHASES			
; FILE REFERENCE: 07678-025001			
; CURRENT FILING DATE: 2001-07-11			
; CURRENT APPLICATION NUMBER: US/09/903,012B			
; PRIOR APPLICATION NUMBER: 60/100,993			
; PRIOR FILING DATE: 1998-09-18			
; PRIOR APPLICATION NUMBER: 60/130,628			
; PRIOR FILING DATE: 1999-04-22			
; PRIOR APPLICATION NUMBER: 60/150,262			
; PRIOR FILING DATE: 1999-08-23			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 56			
; LENGTH: 868			
; TYPE: PRT			
; ORGANISM: Abies grandis			
US-09-903-012B-56			
Query Match 61.2%; Score 2844; DB 4; Length 868;			
Best Local Similarity 62.3%; Pred. No. 2.7e-279;			
Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;			
Qy	12	SLQLSPKVP---PROSTNIIIPFKRSSFQNAHCVRSHLRWN-----CVGIHA 60	
Db	5	SSLSLQIPTAAHHLTANAQSPHFSTTLNAGSSAKRSLYLRWKGKSNKIACVG--E 62	
Qy	61	SAAETRPDLQPOBERFVSR---LNADYHPAVWKDPIDSLTSPNSHATSKSSVDETINKR 117	
Db	63	GGATSVPYQSAEKNDLSSTLVKREFPGFWKDDLLDLSLT--SHKVAAS--DE---KR 115	
Qy	118	IQTLVKEIOCMFOSMGDGETNPSAYDTAVARIPSIDGSGAPQFPOTLOWILNNQLPDGS 177	
Db	116	IETLSEIKNMFRCMVGGETNPSAYDTAVARIPAVDGSNDPHFPETVEWILLQNLKOGS 175	
Qy	178	WGECIFLAYDRVLNTLACLLTLKWNKGDIOVQKGVEFVRKHEMKDEADNHRPSGE 237	
Db	176	WGEGFYFLAYDRILATLACIITLTWRTGETVQKGIEFRTQAGMEDEADSHRPSGE 235	
Qy	238	VVPAMLDEAKSLGLDLPYHLPEISQIHOKRCKLOKIPLVNLHHTQALLYSLLEGQDV 297	
Db	236	IVFPAMLKEAKILGLDLPYDLPELKOIEKREAKUKRIPTDVLIALPTLLYSLLEGQEI 295	
Qy	298	VDQWETNLSQRDGSFLSSPASTACVFMHTQNKRCILHFLNVLFSKFGDYVPCHYPLDLPE 357	
Db	296	VDQKIMKLOSKDGSFLSSPASTAAVFMRTGNKKCLDLFLNFKKFGNHPCHYPLDLPE 355	
Qy	358	RLWAVDTVERLIGDIRYFKKEIKESLDYVTRYWDAERGVGWARCNPIPDVDDTAMGLRIUR 417	
Db	356	RLWAVDTVERLIGDIRHFKBEIKEALDYVYSHWD-ERGIQWARENPVPDIDDTAMGLRIUR 414	
Qy	418	LHGYNVSSDVLNFRDEBKDGFCCFAGQTOIGVTDNLNLYRCQVCPGPKIMEEAKTFTT 477	
Db	415	LHGYNVSSDVLNFKTRDENGEFFCFLGQTOGVTDMLNVNRCVSHVFPGETIMEEAKLCTE 474	
Qy	478	NHLQNALAKNNAPDKWAVKDLPGEVEYAIKYPMHRSMPRLBARSYIEQFGSNDVWLKGT 537	
Db	475	RYLRNALENVDAFDKNAFKKNIRGEVEYALKYPWHKSMRPLBARSYIENYGPDDVWLKGT 534	
Qy	538	VYKMLVYSNEKYLELAKLPNMVQALHQBETHQIHVSWWRESGENDLTFTTRQRPVMEYFSV 597	
Db	535	VYMPYIYSNEKYLELAKLPNFKVQSTHQLDQLRRWKKSSGFTDLNFTTRERVTBIYFSP 594	
Qy	598	AVSMFEPFAACARIAYAKTSCLAVIIDLVDYTHGSLDDLKLFSEAVRWDDISVLSVRDN 657	
Db	595	ASFIPEFSEKREVYKTSNFTVILDDYDAGHSDDLKLFTESEKRWDLSDQDM-PQ 653	
Qy	658	QLKVCFLGLYNTVNGFKDGLKEQGRDLVLYLRKVGWELLASYTEAEWSAAKYVPTFNE 717	
Db	654	QMKICFVGFTNTDIAKGRERQGRDLVLYQNVKVLQLEAYTKEAEWSAAKYVPSFNE 713	
Qy	718	YVENAKVSIATLVNLSIPFTGELLPDVYLQOVDLRSKFLHLVSLTGRLLINDTKTYQAE 777	
Db	714	YIENASVSIALGTWVLSIALFTCEVLTDVLSKIDRESFLQMLGTLGRLLVNDTKTYQAE 773	

Db 475 RYLRNLENDVDAFKWAFKKNIRGEVEYALKYPWHKSMRLEARSYIENYGPDDVWLKGT 534  
Qy 538 VYKMLYVSENEKYLELAKLDENWQALHOKETQHIVSMWRESGFNDLTFTRQRPVEMYFSV 597  
Db 535 VYMPYIISNEKYLELAKLDENWQALHOKETQHIVSMWRESGFNDLTFTRQRPVEMYFSV 594  
Qy 598 AVSMPEPFAACRIAYAKTSCIAVILDDLYDTHGSLDLDLKFSEAVRWDISVLDVRDN 657  
Db 595 ASFIPEPEFSKREYVTKTSNFTVILDDLYDAHGSLDLDLKFSEAVRWDISVLDVRDN 653  
Qy 658 QLKVCFGLYNTVNGFGKDLKEQGRDVLGYLRKYWEGLSYTKAEWSAAKYVPTFNE 717  
Db 654 QMKICFVGFTNFNDIAKEGRERQGRDVLGYLRKYWEGLSYTKAEWSAAKYVPTFNE 713  
Qy 718 YVENAKVSIATATVVLNSIFFTGELLDPYILOQVDLRSKFLHVLSTLGRINDTKTYQAE 777  
Db 714 YIENASVSIALGTVVLISALFTGEVLTDEVLSKIDRESRFLQMLGTGRVNDTKTYQAE 773  
Qy 778 RNRGELVSSVQCYMRENPECTBEALSHVYGIIDNALKELNLANPASNAPLCVRRLLF 837  
Db 774 RQGEVASAIQCYMKDHPKISEEALQHVYSVMENALEELNREFVN--NKIPDIYKRLVF 831  
Qy 838 NTAARVWOLFMYRVDGFGIS--DKEMKDHVSRTLFDPA 873  
Db 832 ETARINQMLFYMQDGLTSLHDMKEIKHVKNCLFQVPA 868  
RESULT 5  
US-09-900-797-56  
; Sequence 56, Application US/09900797  
; Patent No. 6645762  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 66457621, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 56  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Abies grandis  
US-09-900-797-56  
Query Match 61.2%; Score 2844; DB 4; Length 868;  
Best Local Similarity 62.3%; Pred. No. 2.7e-279;  
Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;  
Qy 12 SLQSPKVP---FRQSTNILLPFKRSFGNAQCHVSHLRWLN-----CVGIHA 60  
Db 5 SSSLSQIPTAAHHLTANAQSIPIHFTSTTLNAGSSASKRRSLYLRWKGSKNIIACVG--E 62  
Qy 61 SAAETRPDLQPEERFVSR---LNADYHPAVWKDDFIDSLTSPNSHATSKSSVDETINKR 117  
Db 63 GGATSVIPQSAEKNDLSLSSTLVKEFPFGFKDLDLSLT--SHKVAAS--DE--KR 115  
Qy 118 IOTLVKEIQCMQSQMGDETHPSAYDTAWARIPSIDSGAPQPTLOWILNQLDPS 177  
Db 116 IETLISEIKNFRCMGCTNPSTAYDTAWARIPAVDGSNPHFPETVEWILQQLKDG 175  
Qy 178 WGEICIFLAYDRVLTNLACLTLTKWNKGDITQVQKGEVFRVAKHEEMKDEADNHRPSGFE 237

Db 176 WGEFVFLAYDRILATLACIITLTWRTGTQVQKGEIEFFRTQAGKMEADSHRPSGFE 235  
Qy 238 VYFPMLEDEAKSLGLDLPYHLPTFSIQHOKROKKLOKIPLNVLHNNHQTALLYSLEGLQDV 297  
Db 236 IVFPAMLEKEAKILGLDLPYDLPLKQIIEKREAKKRIPTDVLXALPTTLLYSLEGLQEI 295  
Qy 298 VDMQEIITNLSQSDGSLSSPASTACVFMHTQNKRLHFLNPLVLSKFGDYVPCHPYDLDLFE 357  
Db 296 VDMQKIMKLSQSDGSLSSPASTAAVFMRTGNKKCLDFLNFVLSKFGHNVPCHPYDLDLFE 355  
Qy 358 RLMAVDVTRERIGIDRYPKKEIKESLDYVYRWDAERGVGWACNPIPDVDDDTAMGLRLR 417  
Db 356 RLMAVDVTRERIGIDRHFKKEIKESLDYVYRWDAERGVGWACNPIPDVDDDTAMGLRLR 414  
Qy 418 LHGYNVSSDVLNFRDEKGFPCFAGQTQIGVTDNLNLYRSCVQCFPGKEKIMEAKTFTT 477  
Db 415 LHGYNVSSDVLNFRDENGFEFCFLGQTQVGTDMNVRNCHSVFPGETIMEEAKLCTE 474  
Qy 478 NHLQNALAKNNAFQKWAQKOLPGEVEYAIKYPHRSMPRLSEASYIEQFSGNDVWLKGT 537  
Db 475 RYLRNLENDVDAFKWAFKKNIRGEVEYALKYPWHKSMRLEARSYIENYGPDDVWLKGT 534  
Qy 538 VYKMLYVSENEKYLELAKLDENWQALHOKETQHIVSMWRESGFNDLTFTRQRPVEMYFSV 597  
Db 535 VYMPYIISNEKYLELAKLDENWQALHOKETQHIVSMWRESGFNDLTFTRQRPVEMYFSV 594  
Qy 598 AVSMPEPFAACRIAYAKTSCIAVILDDLYDTHGSLDLDLKFSEAVRWDISVLDVRDN 657  
Db 595 ASFIPEPEFSKREYVTKTSNFTVILDDLYDAHGSLDLDLKFSEAVRWDISVLDVRDN 653  
Qy 658 QLKVCFGLYNTVNGFGKDLKEQGRDVLGYLRKYWEGLSYTKAEWSAAKYVPTFNE 717  
Db 654 QMKICFVGFTNFNDIAKEGRERQGRDVLGYLRKYWEGLSYTKAEWSAAKYVPTFNE 713  
Qy 718 YVENAKVSIATATVVLNSIFFTGELLDPYILOQVDLRSKFLHVLSTLGRINDTKTYQAE 777  
Db 714 YIENASVSIALGTVVLISALFTGEVLTDEVLSKIDRESRFLQMLGTGRVNDTKTYQAE 773  
Qy 778 RNRGELVSSVQCYMRENPECTBEALSHVYGIIDNALKELNLANPASNAPLCVRRLLF 837  
Db 774 RQGEVASAIQCYMKDHPKISEEALQHVYSVMENALEELNREFVN--NKIPDIYKRLVF 831  
Qy 838 NTAARVWOLFMYRVDGFGIS--DKEMKDHVSRTLFDPA 873  
Db 832 ETARINQMLFYMQDGLTSLHDMKEIKHVKNCLFQVPA 868  
RESULT 6  
US-09-234-393-13  
; Sequence 13, Application US/09234393A  
; Patent No. 6265639  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B  
; APPLICANT: Crock, John E  
; APPLICANT: Bohlman, Jorg  
; APPLICANT: Jetter, Reinhard  
; APPLICANT: Steele, Christopher L  
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
; TITLE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: WSUR113345  
; CURRENT APPLICATION NUMBER: US/09/234,393A  
; CURRENT FILING DATE: 1999-01-20  
; EARLIER APPLICATION NUMBER: 60/072,204  
; EARLIER FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 817  
; TYPE: PRT  
; ORGANISM: Abies grandis  
US-09-234-393-13  
Query Match 45.5%; Score 2114.5; DB 3; Length 817;



Db 735 MKDHPECTEALDHIYSILPAVKELTREFLKP-DDVPPACKMLFEETRTVWIFKDG 793

Qy 851 DFGISDKEMKDHVSRTLDFPV 872

Db 794 DFGVSKLEVKDHIKECLIEPL 815

RESULT 8

US-09-865-171-13

; Sequence 13, Application US/09865171

; Patent No. 6451576

; GENERAL INFORMATION:

; APPLICANT: Croteau, Rodney B

; APPLICANT: Crock, John E

; APPLICANT: Bohlmann, Jorg

; APPLICANT: Steele, Christopher L

; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),

; TITLE OF INVENTION: AND METHODS OF USE

; FILE REFERENCE: WSUR117468

; CURRENT APPLICATION NUMBER: US/09/865,171

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: 09/234,393

; PRIOR FILING DATE: 1999-01-20

; PRIOR APPLICATION NUMBER: 60/072,204

; PRIOR FILING DATE: 1998-01-22

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 13

; TYPE: PRT

; LENGTH: 817

; ORGANISM: Abies grandis

US-09-865-171-13

Query Match 45.5%; Score 2114.5; DB 4; Length 817;

Best Local Similarity 49.5%; Pred. No. 3.3e-205;

Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;

Qy 79 RLNADYHPAWKDDFIDSLTSPNSHATSKSSVDETINKRIQTLVKEIQCMFQS--MGDGE 136

Db 25 RRTANPHNPVGVYDLVHSLKSP-----YIDSSYRERAEVLVSEIKAMLPNPAITGDGE 76

Qy 137 T--NPSAYDTAWVARIPSDIGSGAPQPTQLWILNNLPDGSWGECIFLAYDRVLNTL 194

Db 77 SMITPSAYDTAWVARIPSDIGSGAPQPTQLWILNNLPDGSWGECIFLAYDRVLNTL 136

Qy 195 ACLLTKIWNKGDIOVKQGVFVRKMEEMKDEAD--NHRPSGFVVPFAMLDKSLGLD 253

Db 137 SCVLVLKWNVDLQVEQGEFIFKSNLELVKDETDQSLVTDFFIIFPSLLREASRLG 196

Qy 254 LPYHLPFISQIHQKQKQKLOKIPLVNLHNHQTALLSLEGLOVDVWQEIINLQSRDGSF 313

Db 197 LPYDLPYIHLQTKRQRLAKLSREIYAVPSPILLYSLEGIQDIVEMERIMEVQSDGSF 256

Qy 314 LSSPASTACVPMHTQNKRCLEHFLNVLKSGDVVPCVPLDPLERLWAVDTVERLIGDIRY 373

Db 257 LSSPASTACVPMHTGDAKCLFELNSVMIKFNVPCVPLDPLERLWAVDTVERLIGDIRY 316

Qy 374 FKKEIKESLDVYRYWDAERGUGWARCNPIPDVDDTAMGLRILRHGYNVSSVDLENFRD 433

Db 317 FEKEIKEALDVIYRHWN-ERGIWGRNLNPIADLETALGFRLLRLHRYNVSPAFDNFKD 375

Qy 434 EKGDFPFCAGTQGTGVTNMLNLYRCQVCPGEKIMEEAKTFTTNHLQNALAKNNAFDKW 493

Db 376 ANGKFCSTGQFNKDVASMLNLYRASQLAFFGENILDEAKSFATKYUREALEKSETSAW 435

Qy 494 AVKDLGCEVYAIKYPWHRSMPLREARSYTEQFGSNDVWLKGTYYKMLVYSNEKYLELA 553

Db 436 NKNQLNSEIKYALKTSWASVPRVEAKYQCVYRPDYARIKACVYKLPYVYNEKEFLELG 495

Qy 554 KLDENMVALHKEQTHQIVSWRRSGFNDLFTFQRPVEMTFSVAVSNFEPEFAACRIAY 613

Db 496 KLDENIIQSIHQEEMKNVTSWFRDGLPLFTFARERPLEFYFLVAAGTYEPQYAKRCLPLF 555

Qy 614 AKTISCLAVILDDYDTHGSLDDKLSEAVRRWDISVLSVRNQLKVCFLGLYNTVNGF 673

Db 556 TKVACLQTVLDDMYDTYGLDELKLFTEAVRRWDLSTENLPD-YMKLCYQIYYDIVEH 614

Qy 674 GKGLKEQGRDVLGRLKWKVEGLLASVYTKAEWSAAKYVTFEYENAKVSIALATVVL 733

Db 615 AWEAKEQGRRELVSFFRKGWEDYLLGYEEAEWLAAYVFTLDEYIKNGITSIGORILL 674

Qy 734 NSI-FFTCELLPDYILQOVDL--RSKFLHLVSLTGRLLINDTKTYOARNRGELVSSVOCY 790

Db 675 SGVLMDGQLLSQEALEKVDYPCRRVLTSLNLSRLADDTKTYAKAEKARGELASSIECY 734

Qy 791 MRENPECTEALSHVYGIIDNALKELNWLANPASNAPLCVRELLFNFTFARVMOLFVYMR 850

Db 735 MKDHPECTEALDHIYSILPAVKELTREFLKP-DDVPPACKMLFEETRTVWIFKDG 793

Qy 851 DFGISDKEMKDHVSRTLDFPV 872

Db 794 DFGVSKLEVKDHIKECLIEPL 815

RESULT 9

US-09-865-171-40

; Sequence 40, Application US/09865171

; Patent No. 6451576

; GENERAL INFORMATION:

; APPLICANT: Croteau, Rodney B

; APPLICANT: Crock, John E

; APPLICANT: Bohlmann, Jorg

; APPLICANT: Steele, Christopher L

; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),

; TITLE OF INVENTION: AND METHODS OF USE

; FILE REFERENCE: WSUR117468

; CURRENT APPLICATION NUMBER: US/09/865,171

; CURRENT FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: 09/234,393

; PRIOR FILING DATE: 1999-01-20

; PRIOR APPLICATION NUMBER: 60/072,204

; PRIOR FILING DATE: 1998-01-22

; NUMBER OF SEQ ID NOS: 55

; SOFTWARE: Patent In Ver. 2.0

; SEQ ID NO 40

; TYPE: PRT

; LENGTH: 817

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: protein

; NAME/KEY: misc.feature

; LOCATION: (1)-(817)

; OTHER INFORMATION: Computer-generated protein sequence

US-09-865-171-40

Query Match 45.5%; Score 2114.5; DB 4; Length 817;

Best Local Similarity 49.5%; Pred. No. 3.3e-205;

Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;

Qy 79 RLNADYHPAWKDDFIDSLTSPNSHATSKSSVDETINKRIQTLVKEIQCMFQS--MGDGE 136

Db 25 RRTANPHNPVGVYDLVHSLKSP-----YIDSSYRERAEVLVSEIKAMLPNPAITGDGE 76

Qy 137 T--NPSAYDTAWVARIPSDIGSGAPQPTQLWILNNLPDGSWGECIFLAYDRVLNTL 194

Db 77 SMITPSAYDTAWVARIPSDIGSGAPQPTQLWILNNLPDGSWGECIFLAYDRVLNTL 136

Qy 195 ACLLTKIWNKGDIOVKQGVFVRKMEEMKDEAD--NHRPSGFVVPFAMLDKSLGLD 253

Db 137 SCVLVLKWNVDLQVEQGEFIFKSNLELVKDETDQSLVTDFFIIFPSLLREASRLG 196

Qy 254 LPYHLPFISQIHQKQKQKLOKIPLVNLHNHQTALLSLEGLOVDVWQEIINLQSRDGSF 313

Db 197 LPYDLPYIHLQTKRQRLAKLSREIYAVPSPILLYSLEGIQDIVEMERIMEVQSDGSF 256

Qy 314 LSSPASTACVFMHTONKRCLEHFLNPLSKFGDYVPCVPLDLFERLWAVDTVERLGDY 373  
 Db 257 LSSPASTACVFMHTONKRCLEHFLNPLSKFGDYVPCVPLDLFERLWAVDTVERLGDY 316  
 Qy 374 FKKEIKESLDVYRYWDAERGVMGWCNPIPDVDDTAMGLRIILRHGYNVSSDVLNFRD 433  
 Db 317 FKKEIKESLDVYRYWDAERGVMGWCNPIPDVDDTAMGLRIILRHGYNVSSDVLNFRD 375  
 Qy 434 EKGDFPFCAGOTQIGVTNINLYRCSQVCPGKEIMEEAKTFTTNHLONALAKNAPDKW 493  
 Db 376 ANGKFCSTGOFNKDVASMLNLYRASQVCPGKEIMEEAKTFTTNHLONALAKNAPDKW 435  
 Qy 494 AVKDLPEVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTYVYKMLYVSNKYLELA 553  
 Db 436 NNKQNLSEIYKALKTSWHASVPRVEAKRYCQVYRDPYARIKACVYKLPYVNNKFLPL 495  
 Qy 554 KLDNFMVQALHOKETOHIIVSWMRSGFNDLFTTRQRPVEMYFVSAVSMFEPEFAACRIAY 613  
 Db 496 KLDNFMVQALHOKETOHIIVSWMRSGFNDLFTTRQRPVEMYFVSAVSMFEPEFAACRIAY 555  
 Qy 614 AKTSCLAVILDDLYDTHGSLDDKLPSEAVRRWDISVLDSVRDNLQKVCFLGLYNTVNGF 673  
 Db 556 TKVACLQVLDMDYDTGLDELKLPTEAVRRWDLSTENLPD-YMKLCYQIYYDIVEH 614  
 Qy 674 GKGLKEQGRDVLGYLRKWEGLLASVYKAEWSAAKYVPTFNEVENAKVSIATVVL 733  
 Db 615 AWEAKEQGRDVLGYLRKWEGLLASVYKAEWSAAKYVPTFNEVENAKVSIATVVL 733  
 Qy 734 NSI-PFTGELLPDYILQOVDL--RSKFLHLVSLTGLRLINDTKTYOABNRGELVSSVQCY 790  
 Db 675 SGVLMDQQLLSQALEKVDYVGRVLTSLNLSRLADDTKYKAEKARGELASSIECY 734  
 Qy 791 MRENPECTEERALSHTVYGIIDNALKELNWLNPASNAPLCVRRLLFNTARVMOLFMYR 850  
 Db 735 MKDHPCTEERALSHTVYGIIDNALKELNWLNPASNAPLCVRRLLFNTARVMOLFMYR 793  
 Qy 851 DGFISDKEMKDHVSRITLFDPV 872  
 Db 794 DGFISDKEMKDHVSRITLFDPV 872

RESULT 10  
 US-09-234-393-38  
 ; Sequence 38, Application US/09234393A  
 ; Patent No. 6265639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney B  
 ; APPLICANT: Crock, John E  
 ; APPLICANT: Bohlman, Jorg  
 ; APPLICANT: Jetter, Reinhard  
 ; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
 ; TITLE OF INVENTION: AND METHODS OF USE  
 ; FILE REFERENCE: WSUR113345  
 ; CURRENT APPLICATION NUMBER: US/09/234,393A  
 ; CURRENT FILING DATE: 1999-01-20  
 ; EARLIER APPLICATION NUMBER: 60/072,204  
 ; EARLIER FILING DATE: 1998-01-22  
 ; NUMBER OF SEQ ID NOS: 55  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 38  
 ; LENGTH: 817  
 ; TYPE: PRT  
 ; ORGANISM: Abies grandis  
 US-09-234-393-38

Query Match 45.5%; Score 2113.5; DB 3; Length 817;  
 Best Local Similarity 49.5%; Pred. No. 4.2e-205;  
 Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;  
 Qy 79 RLNDYHPAVKDDPIDSITSPNSHATSKSSVDEITNKRIQTLVKEIQCMFQS--MGDGE 136  
 Db 25 RTANPHNVWGYDVLVHLSKSP-----YIDSSYRRAEVLVSEIKVMLNPAITGDGE 76

Qy 137 T--NPSAYDTAWARIPSIDSGNAPQPPOTLOWILNANLPDGSWGEBECIFLAYDVLNLT 194  
 Db 77 SMITPSAYDTAWARIPSIDSGNAPQPPOTLOWILNANLPDGSWGEBECIFLAYDVLNLT 136  
 Qy 195 ACILTLTKMNKGDITQVKGVEFVRKHEEMKDEAD-NHRPSGFEVFPFAMLDKAKSLGLD 253  
 Db 137 SCVLVLLKNNVGLDQVEQIEFKSNLELVKDETDQDLSVDFEIIIPSLRLAQAQLRG 196  
 Qy 254 LPYHLPISQIHKRQKQLOKIPLVNLHNTQALLYSLEGLODVVDWQEIETNLOSQSGSP 313  
 Db 197 LPYDLPYIHLQTKRQERLAKLSREIYAVPSPLLYSLEGIQDIVEIMERIMEVQSDGFS 256  
 Qy 314 LSSPASTACVFMHTONKRCLEHFLNPLSKFGDYVPCVPLDLFERLWAVDTVERLGDY 373  
 Db 257 LSSPASTACVFMHTONKRCLEHFLNPLSKFGDYVPCVPLDLFERLWAVDTVERLGDY 316  
 Qy 374 FKKEIKESLDVYRYWDAERGVMGWCNPIPDVDDTAMGLRIILRHGYNVSSDVLNFRD 433  
 Db 317 FKKEIKESLDVYRYWDAERGVMGWCNPIPDVDDTAMGLRIILRHGYNVSSDVLNFRD 375  
 Qy 434 EKGDFPFCAGOTQIGVTNINLYRCSQVCPGKEIMEEAKTFTTNHLONALAKNAPDKW 493  
 Db 376 ANGKFCSTGOFNKDVASMLNLYRASQVCPGKEIMEEAKTFTTNHLONALAKNAPDKW 435  
 Qy 494 AVKDLPEVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTYVYKMLYVSNKYLELA 553  
 Db 436 NNKQNLSEIYKALKTSWHASVPRVEAKRYCQVYRDPYARIKACVYKLPYVNNKFLPL 495  
 Qy 554 KLDNFMVQALHOKETOHIIVSWMRSGFNDLFTTRQRPVEMYFVSAVSMFEPEFAACRIAY 613  
 Db 496 KLDNFMVQALHOKETOHIIVSWMRSGFNDLFTTRQRPVEMYFVSAVSMFEPEFAACRIAY 555  
 Qy 614 AKTSCLAVILDDLYDTHGSLDDKLPSEAVRRWDISVLDSVRDNLQKVCFLGLYNTVNGF 673  
 Db 556 TKVACLQVLDMDYDTGLDELKLPTEAVRRWDLSTENLPD-YMKLCYQIYYDIVEH 614  
 Qy 674 GKGLKEQGRDVLGYLRKWEGLLASVYKAEWSAAKYVPTFNEVENAKVSIATVVL 733  
 Db 615 AWEAKEQGRDVLGYLRKWEGLLASVYKAEWSAAKYVPTFNEVENAKVSIATVVL 733  
 Qy 734 NSI-PFTGELLPDYILQOVDL--RSKFLHLVSLTGLRLINDTKTYOABNRGELVSSVQCY 790  
 Db 675 SGVLMDQQLLSQALEKVDYVGRVLTSLNLSRLADDTKYKAEKARGELASSIECY 734  
 Qy 791 MRENPECTEERALSHTVYGIIDNALKELNWLNPASNAPLCVRRLLFNTARVMOLFMYR 850  
 Db 735 MKDHPCTEERALSHTVYGIIDNALKELNWLNPASNAPLCVRRLLFNTARVMOLFMYR 793  
 Qy 851 DGFISDKEMKDHVSRITLFDPV 872  
 Db 794 DGFISDKEMKDHVSRITLFDPV 872

RESULT 11  
 US-09-234-393-42  
 ; Sequence 42, Application US/09234393A  
 ; Patent No. 6265639  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Croteau, Rodney B  
 ; APPLICANT: Crock, John E  
 ; APPLICANT: Bohlman, Jorg  
 ; APPLICANT: Jetter, Reinhard  
 ; APPLICANT: Steele, Christopher L  
 ; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
 ; TITLE OF INVENTION: AND METHODS OF USE  
 ; FILE REFERENCE: WSUR113345  
 ; CURRENT APPLICATION NUMBER: US/09/234,393A  
 ; CURRENT FILING DATE: 1999-01-20  
 ; EARLIER APPLICATION NUMBER: 60/072,204  
 ; EARLIER FILING DATE: 1998-01-22  
 ; NUMBER OF SEQ ID NOS: 55  
 ; SOFTWARE: Patent In Ver. 2.0







US-09-398-395A-46

Query Match 44.9% Score 2085.5; DB 4; Length 782;  
Best Local Similarity 49.6%; Pred. No. 2.7e-202;  
Matches 391; Conservative 158; Mismatches 221; Indels 19; Gaps 9;

QY 92 DFIDSLTSPNHSATSKSSVDETINKRIQTIVKETOQMFQS--MGDET--NPSAYDTAWV 147  
Db 3 DLVHSLKSP-----YIDSSYREAEVLVSEIKVMLNPAITGDGSMITPFSAYDTAWV 54  
QY 148 ARIPSIDSGAPQPPQILOWLNNQLPDGSNGECIFLAYDRLVNTLACLTLLKWNKGD 207  
Db 55 ARVPAIDGSAEPQPPQILOWLNNQLPDGSNGECIFLAYDRLVNTLACLTLLKWNKGD 114  
QY 208 IQOVGVFVFRKWEEMKDEAD--NHRPSGFVFPFAMLDKSLGLDLPYHLPFISQIHQ 266  
Db 115 LQVSGGIEFISNLELVKDETDQSLVTDFFIIFPSLLREAAQSLRLGLPYLPIHLLQ 174  
QY 267 KRQKLOKIPLNVLNHNHOTALLYSLLEGQDVVDQWQEIITNLSRDGSLSSPASTACVFMH 326  
Db 175 KRQERLAKLSREIYAVSPPLYSLEGIQDIVERIMEVQSDGSLSSPASTACVFMH 234  
QY 327 TQNKCEHLFNLVSKFGDYVPCYPLDLFERLMAVDTVRIGIDRYPKKEIKESLDVYV 386  
Db 235 TGDACLEFLNSVMIKFGNFVPCLYPVDLLERLLIVDNIYRIGIYRHFKEIKEALDYYV 294  
QY 387 RYWDARGVGWARNCPIDVDVDTAMGLRILRLHGVNVSVDVLENPRDEKGFPCFAGQTO 446  
Db 295 RHNW--ERGIGWGLNPIADLETTALGFRLLRLHRYNVPSPAIQDNFKDANGKFCISGTQPN 353  
QY 447 IGVTDNLNLYRCSQVCFPEGEKIMEEAKFTTNHLQNALAKNNAFQKWAQKQDLPGEVEYA 506  
Db 354 KDVASMLNLYRASQALAFGENILDEAKSFATKYLREALKESSTSSANNKQNLQSEIKYA 413  
QY 507 IKYPHRSMPRLARSYIEQFGSNDVWLKTYVYKMLYVSNKYLELAKLDFNNVQALHOK 566  
Db 414 LKTSWASVPRVEAKRYCQVYRDPYARIKACVYKLPYVYNNKFLGKLDNFNIIOSIHQ 473  
QY 567 ETQHIVSWREGSFNDLTTRORPVMYFSAVSMFEPFAACRIAYAKTSCIALVILDDL 626  
Db 474 EMKNVTSWFRDGLPLFTFARERPLEFVFLVAAGTYEPQYAKRFLTKVACLTQVILDDM 533  
QY 627 YDTHGSLDDLKLFSEAVRWDLISVLDSVRDNLKVCFLGLYNTVNGFGKDGKEQGRDVL 686  
Db 534 YDTYGTLDLKLFTFAVRWDLISFTENLPD--YMKLCYQIYYDIDVHEVAWEAKEQGRELV 592  
QY 687 GYLKRWEGILLASYTKEAESAAYVPTFNEVYENAKVSIATATVVLNSI--FFTGLLPPD 745  
Db 593 SFRKGMEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGQRILLSGVLIMDGQLLSQ 652  
QY 746 YILQVVDL--RSKFLHLVSLTGLINDTKTYQAERNRNGELVSSVQCYMRENPECTEERAL 803  
Db 653 EALEKVDYFGRVLTTELNSLSRLADDTKTYKAEKARGELASSIECYMKDHPCTEERAL 712  
QY 804 SHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTRVLMQVYRDGFGISDKEMKOH 863  
Db 713 DHIYSILEPAVKELTREFLKP--DDVPFACKKMLFEETRVTWVIFKODGDFGVSKLEVKOH 771  
QY 864 VSRTLDPV 872  
Db 772 IKECLIEPL 780

Search completed: August 17, 2004, 21:03:48  
Job time : 21.8728 secs

us-10-041-007-2.ra1

Db 115 LQVEQIEBFIKSNELVXDQDQSLVTDFFIIFPSLLREAAQSLRLGLPYDLPYIHLIQT 174  
QY 267 KRQKLOKIPLNVLNHNHOTALLYSLLEGQDVVDQWQEIITNLSRDGSLSSPASTACVFMH 326  
Db 175 KRQERLAKLSREIYAVSPPLYSLEGIQDIVERIMEVQSDGSLSSPASTACVFMH 234  
QY 327 TQNKCEHLFNLVSKFGDYVPCYPLDLFERLMAVDTVRIGIDRYPKKEIKESLDVYV 386  
Db 235 TGDACLEFLNSVMIKFGNFVPCLYPVDLLERLLIVDNIYRIGIYRHFKEIKEALDYYV 294  
QY 387 RYWDARGVGWARNCPIDVDVDTAMGLRILRLHGVNVSVDVLENPRDEKGFPCFAGQTO 446  
Db 295 RHNW--ERGIGWGLNPIADLETTALGFRLLRLHRYNVPSPAIQDNFKDANGKFCISGTQPN 353  
QY 447 IGVTDNLNLYRCSQVCFPEGEKIMEEAKFTTNHLQNALAKNNAFQKWAQKQDLPGEVEYA 506  
Db 354 KDVASMLNLYRASQALAFGENILDEAKSFATKYLREALKESSTSSANNKQNLQSEIKYA 413  
QY 507 IKYPHRSMPRLARSYIEQFGSNDVWLKTYVYKMLYVSNKYLELAKLDFNNVQALHOK 566  
Db 414 LKTSWASVPRVEAKRYCQVYRDPYARIKACVYKLPYVYNNKFLGKLDNFNIIOSIHQ 473  
QY 567 ETQHIVSWREGSFNDLTTRORPVMYFSAVSMFEPFAACRIAYAKTSCIALVILDDL 626  
Db 474 EMKNVTSWFRDGLPLFTFARERPLEFVFLVAAGTYEPQYAKRFLTKVACLTQVILDDM 533  
QY 627 YDTHGSLDDLKLFSEAVRWDLISVLDSVRDNLKVCFLGLYNTVNGFGKDGKEQGRDVL 686  
Db 534 YDTYGTLDLKLFTFAVRWDLISFTENLPD--YMKLCYQIYYDIDVHEVAWEAKEQGRELV 592  
QY 687 GYLKRWEGILLASYTKEAESAAYVPTFNEVYENAKVSIATATVVLNSI--FFTGLLPPD 745  
Db 593 SFRKGMEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGQRILLSGVLIMDGQLLSQ 652  
QY 746 YILQVVDL--RSKFLHLVSLTGLINDTKTYQAERNRNGELVSSVQCYMRENPECTEERAL 803  
Db 653 EALEKVDYFGRVLTTELNSLSRLADDTKTYKAEKARGELASSIECYMKDHPCTEERAL 712  
QY 804 SHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTRVLMQVYRDGFGISDKEMKOH 863  
Db 713 DHIYSILEPAVKELTREFLKP--DDVPFACKKMLFEETRVTWVIFKODGDFGVSKLEVKOH 771  
QY 864 VSRTLDPV 872  
Db 772 IKECLIEPL 780

RESULT 15  
US-09-398-395A-46  
; Sequence 46, Application US/09398395A  
; Patent No. 6468772  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 64687721, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/398,395A  
; CURRENT FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 46  
; LENGTH: 782  
; TYPE: PRT  
; ORGANISM: Abies grandis

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 21:02:37 ; Search time 47.4515 Seconds  
(without alignments)  
5775.546 Million cell updates/sec

Title: US-10-041-007-2  
Perfect score: 4645  
Sequence: 1 MAGVLFANLPCSLQLSPKVP.....GISDKEMKONVSRITLDPVA 873

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
4: /cgn2\_6/ptodata/1/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	4645	100.0	873	12	US-10-041-018-398
2	4645	100.0	873	13	US-10-041-007-2
3	4645	100.0	873	13	US-10-041-007-33
4	4319	93.0	814	13	US-10-041-007-35
5	4224	90.9	795	13	US-10-041-007-37
6	3972	85.5	746	13	US-10-041-007-39
7	2844	61.2	868	9	US-09-887-586A-56
8	2844	61.2	868	9	US-09-903-012-56
9	2844	61.2	868	10	US-09-900-797-56
10	2844	61.2	868	12	US-09-893-820-56
11	2844	61.2	868	12	US-10-041-018-388
12	2844	61.2	868	13	US-10-041-007-4
13	2114.5	45.5	817	13	US-10-041-007-15
14	2085.5	44.9	782	9	US-09-887-586A-46
15	2085.5	44.9	782	9	US-09-903-012-46

16	2085.5	44.9	782	10	US-09-900-797-46	Sequence 46, Appl
17	2085.5	44.9	782	12	US-09-893-820-46	Sequence 46, Appl
18	2085.5	44.9	782	13	US-10-041-007-16	Sequence 16, Appl
19	2085.5	44.9	782	14	US-10-025-145A-16	Sequence 16, Appl
20	1959	42.2	862	9	US-09-887-586A-44	Sequence 44, Appl
21	1959	42.2	862	9	US-09-903-012-44	Sequence 44, Appl
22	1959	42.2	862	10	US-09-900-797-44	Sequence 44, Appl
23	1959	42.2	862	12	US-09-893-820-44	Sequence 44, Appl
24	1959	42.2	862	12	US-10-041-018-386	Sequence 386, App
25	1959	42.2	862	13	US-10-041-007-41	Sequence 41, Appl
26	1411	30.4	787	12	US-10-041-018-385	Sequence 385, App
27	1377.5	29.7	815	9	US-09-371-307-88	Sequence 88, Appl
28	1377.5	29.7	815	15	US-10-401-321-88	Sequence 88, Appl
29	1318	28.4	823	12	US-10-041-018-396	Sequence 396, App
30	1311	28.2	840	16	US-10-437-963-182585	Sequence 182585,
31	1303	28.1	801	12	US-10-041-018-394	Sequence 394, App
32	1252	27.0	581	9	US-09-887-586A-48	Sequence 48, Appl
33	1252	27.0	581	9	US-09-903-012-48	Sequence 48, Appl
34	1252	27.0	581	10	US-09-900-797-48	Sequence 48, Appl
35	1252	27.0	581	12	US-09-893-820-48	Sequence 48, Appl
36	1252	27.0	581	13	US-10-041-007-18	Sequence 18, Appl
37	1215.5	26.2	630	14	US-10-025-145A-67	Sequence 67, Appl
38	1206	26.0	577	14	US-10-025-145A-18	Sequence 18, Appl
39	1194.5	25.7	637	9	US-09-887-586A-58	Sequence 58, Appl
40	1194.5	25.7	637	9	US-09-903-012-58	Sequence 58, Appl
41	1194.5	25.7	637	10	US-09-900-797-58	Sequence 58, Appl
42	1194.5	25.7	637	12	US-09-893-820-58	Sequence 58, Appl
43	1194.5	25.7	637	13	US-10-041-007-24	Sequence 24, Appl
44	1194.5	25.7	637	13	US-10-025-145A-6	Sequence 6, Appl
45	1193	25.7	637	13	US-10-041-007-28	Sequence 28, Appl

## ALIGNMENTS

### RESULT 1

US-10-041-018-398  
; Sequence 398, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuoka, Seiichi P.T.  
; APPLICANT: Hart, Elizabeth A.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 398  
; LENGTH: 873  
; TYPE: PRT  
; ORGANISM: Ginkgo biloba  
US-10-041-018-398

Query Match 100.0%; Score 4645; DB 12; Length 873;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAGVLFANLPCSLQLSPKVPFRQSTNTLIPFKHSSFGFNAQHCVRSHLRWNCVGIHA	60
Db	1	MAGVLFANLPCSLQLSPKVPFRQSTNTLIPFKHSSFGFNAQHCVRSHLRWNCVGIHA	60
Qy	61	SAATRPDQLPQERFVSRNLNADYHPAVKDDFIDSLTSPNSHATSKSSVDETINKRIQT	120
Db	61	SAATRPDQLPQERFVSRNLNADYHPAVKDDFIDSLTSPNSHATSKSSVDETINKRIQT	120
Qy	121	LVKEIQCMFQSMGDETNPSAYDTAWARIPSIDSGAPQFPQTQLWLNQLPDGSGWGE	180
Db	121	LVKEIQCMFQSMGDETNPSAYDTAWARIPSIDSGAPQFPQTQLWLNQLPDGSGWGE	180
Qy	181	ECIFLAYDRVLTILACLLTLKIWNKGDIQVQKGVFVRKMEEMKDEADNHRPSGFVVF	240



; PRIOR FILING DATE: 2001-01-05  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 33  
 ; LENGTH: 873  
 ; TYPE: PR1  
 ; ORGANISM: Ginkgo biloba  
 US-10-041-007-33

Query Match 100.0%; Score 4645; DB 13; Length 873;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAGVLPANLPCSLQSPKVPVPROSTNIIIPHKRSSFGFNAQHCVRSHLRNRWCVGIIHA	60
Db	1	MAGVLPANLPCSLQSPKVPVPROSTNIIIPHKRSSFGFNAQHCVRSHLRNRWCVGIIHA	60
Qy	61	SAAEIRPDQLPOEERFVSRLNADYHPAVWKDDFIDSLTSPNSHATSKSSVDETINKRIQT	120
Db	61	SAAEIRPDQLPOEERFVSRLNADYHPAVWKDDFIDSLTSPNSHATSKSSVDETINKRIQT	120
Qy	121	LVKEIQCMFQSMGDETNPSAYDTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGE	180
Db	121	LVKEIQCMFQSMGDETNPSAYDTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGE	180
Qy	181	ECIFLAYDRVLNTLACLLTLKWNKGDIOVQKGVFVRKHEEMKDEADNHRPSGFEVVF	240
Db	181	ECIFLAYDRVLNTLACLLTLKWNKGDIOVQKGVFVRKHEEMKDEADNHRPSGFEVVF	240
Qy	241	PAMLDEAKSLGLDLPYHLPFISQIHQRKQKLOKIPLVNLNHQTALLYSLEGLQDVVDW	300
Db	241	PAMLDEAKSLGLDLPYHLPFISQIHQRKQKLOKIPLVNLNHQTALLYSLEGLQDVVDW	300
Qy	301	QEITNLQSRDGSFLSSPASTACVFMHTQNKRCFLHFLNVLKSGDYVPCHPYPLDLFERLW	360
Db	301	QEITNLQSRDGSFLSSPASTACVFMHTQNKRCFLHFLNVLKSGDYVPCHPYPLDLFERLW	360
Qy	361	AVDTVERLIGIDRYFKKEIKESLDYVRYWDAERGVGWARCNPIPDVDDTAMGLRLHLHG	420
Db	361	AVDTVERLIGIDRYFKKEIKESLDYVRYWDAERGVGWARCNPIPDVDDTAMGLRLHLHG	420
Qy	421	YVSSDVLNFRDEKGDFFCFAGQTQIGVTDNLNLRCQVCFPGKEIMEEAKFTTNHL	480
Db	421	YVSSDVLNFRDEKGDFFCFAGQTQIGVTDNLNLRCQVCFPGKEIMEEAKFTTNHL	480
Qy	481	QNALAKNNAFDKWA VKKDLPGVEYAIKYPHRSMPLEARSYIEQFGSNDVWLKGTYYK	540
Db	481	QNALAKNNAFDKWA VKKDLPGVEYAIKYPHRSMPLEARSYIEQFGSNDVWLKGTYYK	540
Qy	541	MLYVSNKYLELAKLDFNMVQALHOKETQHIIVSWWRSGFNDLTFTTRQRPVEMTFSVAVS	600
Db	541	MLYVSNKYLELAKLDFNMVQALHOKETQHIIVSWWRSGFNDLTFTTRQRPVEMTFSVAVS	600
Qy	601	MPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLK	660
Db	601	MPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLK	660
Qy	661	VCFLGLYNTVNGFGDKGKEQGRDVLGYLRKWEGLLASYYTKEAWEWSAAKVTPTFNEIVE	720
Db	661	VCFLGLYNTVNGFGDKGKEQGRDVLGYLRKWEGLLASYYTKEAWEWSAAKVTPTFNEIVE	720
Qy	721	NAKVSIALATVVLNSIIFTGELLPPYIILQQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNR	780
Db	721	NAKVSIALATVVLNSIIFTGELLPPYIILQQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNR	780
Qy	781	GELVSSVQCYMRENPECTEESALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNFA	840
Db	781	GELVSSVQCYMRENPECTEESALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNFA	840
Qy	841	RVNQLPYMYRDGFGISDKEMKDHVSRTLFDPVA	873
Db	841	RVNQLPYMYRDGFGISDKEMKDHVSRTLFDPVA	873

RESULT 4

US-10-041-007-35  
 ; Sequence 35, Application US/10041007  
 ; Publication No. US20020164736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matauda, Seiichi P.T.  
 ; APPLICANT: Schepmann, Hala G.  
 ; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
 ; FILE REFERENCE: P02081US1  
 ; CURRENT APPLICATION NUMBER: US/10/041.007  
 ; CURRENT FILING DATE: 2002-01-07  
 ; PRIOR APPLICATION NUMBER: US 60/259,881  
 ; PRIOR FILING DATE: 2001-01-05  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 35  
 ; LENGTH: 814  
 ; TYPE: PR1  
 ; ORGANISM: Ginkgo biloba  
 US-10-041-007-35

Query Match 93.0%; Score 4319; DB 13; Length 814;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	61	SAAEIRPDQLPOEERFVSRLNADYHPAVWKDDFIDSLTSPNSHATSKSSVDETINKRIQT	120
Db	2	SAAEIRPDQLPOEERFVSRLNADYHPAVWKDDFIDSLTSPNSHATSKSSVDETINKRIQT	61
Qy	121	LVKEIQCMFQSMGDETNPSAYDTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGE	180
Db	62	LVKEIQCMFQSMGDETNPSAYDTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGE	121
Qy	181	ECIFLAYDRVLNTLACLLTLKWNKGDIOVQKGVFVRKHEEMKDEADNHRPSGFEVVF	240
Db	122	ECIFLAYDRVLNTLACLLTLKWNKGDIOVQKGVFVRKHEEMKDEADNHRPSGFEVVF	181
Qy	241	PAMLDEAKSLGLDLPYHLPFISQIHQRKQKLOKIPLVNLNHQTALLYSLEGLQDVVDW	300
Db	182	PAMLDEAKSLGLDLPYHLPFISQIHQRKQKLOKIPLVNLNHQTALLYSLEGLQDVVDW	241
Qy	301	QEITNLQSRDGSFLSSPASTACVFMHTQNKRCFLHFLNVLKSGDYVPCHPYPLDLFERLW	360
Db	242	QEITNLQSRDGSFLSSPASTACVFMHTQNKRCFLHFLNVLKSGDYVPCHPYPLDLFERLW	301
Qy	361	AVDTVERLIGIDRYFKKEIKESLDYVRYWDAERGVGWARCNPIPDVDDTAMGLRLHLHG	420
Db	302	AVDTVERLIGIDRYFKKEIKESLDYVRYWDAERGVGWARCNPIPDVDDTAMGLRLHLHG	361
Qy	421	YVSSDVLNFRDEKGDFFCFAGQTQIGVTDNLNLRCQVCFPGKEIMEEAKFTTNHL	480
Db	362	YVSSDVLNFRDEKGDFFCFAGQTQIGVTDNLNLRCQVCFPGKEIMEEAKFTTNHL	421
Qy	481	QNALAKNNAFDKWA VKKDLPGVEYAIKYPHRSMPLEARSYIEQFGSNDVWLKGTYYK	540
Db	422	QNALAKNNAFDKWA VKKDLPGVEYAIKYPHRSMPLEARSYIEQFGSNDVWLKGTYYK	481
Qy	541	MLYVSNKYLELAKLDFNMVQALHOKETQHIIVSWWRSGFNDLTFTTRQRPVEMTFSVAVS	600
Db	482	MLYVSNKYLELAKLDFNMVQALHOKETQHIIVSWWRSGFNDLTFTTRQRPVEMTFSVAVS	541
Qy	601	MPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLK	660
Db	542	MPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLK	601
Qy	661	VCFLGLYNTVNGFGDKGKEQGRDVLGYLRKWEGLLASYYTKEAWEWSAAKVTPTFNEIVE	720
Db	602	VCFLGLYNTVNGFGDKGKEQGRDVLGYLRKWEGLLASYYTKEAWEWSAAKVTPTFNEIVE	661
Qy	721	NAKVSIALATVVLNSIIFTGELLPPYIILQQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNR	780
Db	662	NAKVSIALATVVLNSIIFTGELLPPYIILQQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNR	721

Db 542 AVILDDLTDHSGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLYNTVNGFGKGLK 601  
Qy 680 EQGSDVLGYLRKWEGLGLASVTKEAWSAKYYPTFNEYVENAKVSIATATVVLNSIPFT 739  
Db 602 EQGSDVLGYLRKWEGLGLASVTKEAWSAKYYPTFNEYVENAKVSIATATVVLNSIPFT 661  
Qy 740 GELLDPYILQQVDLRSKPLHLVSLTGRINDTKTYQAERNRNGELVSSVQCYMRENPECTE 799  
Db 662 GELLDPYILQQVDLRSKPLHLVSLTGRINDTKTYQAERNRNGELVSSVQCYMRENPECTE 721  
Qy 800 BEALSHVYGIIDNALKELNWEANPASNAPLCVRRLLFNTARVMQLFMYRDGFGISDKE 859  
Db 722 BEALSHVYGIIDNALKELNWEANPASNAPLCVRRLLFNTARVMQLFMYRDGFGISDKE 781  
Qy 860 MKDHVSRTLPDPA 873  
Db 782 MKDHVSRTLPDPA 795  
RESULT 6  
US-10-041-007-39  
; Sequence 39, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 39  
; LENGTH: 746  
; TYPE: PRT  
; ORGANISM: Ginkgo biloba  
US-10-041-007-39  
Query Match 85.5%; Score 3972; DB 13; Length 746;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 128 MFQSMGDTNPSAYDTAWARIPSIDGSGAPQPTQLQWLNQPLDGSWGSECFILAY 187  
Db 1 MFQSMGDTNPSAYDTAWARIPSIDGSGAPQPTQLQWLNQPLDGSWGSECFILAY 60  
Qy 188 DRVLNTLACLLTLKIWNKGDIOVQKGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEA 247  
Db 61 DRVLNTLACLLTLKIWNKGDIOVQKGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEA 120  
Qy 248 KSLGLDLPYHLPPISQIHQKQKLOKIPLVNLHNTALLYSLGLQDVVDWQEIITNLQ 307  
Db 121 KSLGLDLPYHLPPISQIHQKQKLOKIPLVNLHNTALLYSLGLQDVVDWQEIITNLQ 180  
Qy 308 SRDGSFLSSPASTACVFMHTQNKRCCLHFLNVLKFGDYVPCHPYPLDLFRLWADTVR 367  
Db 181 SRDGSFLSSPASTACVFMHTQNKRCCLHFLNVLKFGDYVPCHPYPLDLFRLWADTVR 240  
Qy 368 LGIDRYPKKEIKESLDYVYRYWDAERGVCWACNPIPDVDDTAMGLRILRLHGYNVSSDV 427  
Db 241 LGIDRYPKKEIKESLDYVYRYWDAERGVCWACNPIPDVDDTAMGLRILRLHGYNVSSDV 300  
Qy 428 LENFRDEKGFPCFAGOTQIGVTDNLNLYRCSQVCFPEKIMEEAKFTTTHLQNALAKN 487  
Db 301 LENFRDEKGFPCFAGOTQIGVTDNLNLYRCSQVCFPEKIMEEAKFTTTHLQNALAKN 360  
Qy 488 NAFDKWAKKDLPGVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTYKMLYVSN 547  
Db 361 NAFDKWAKKDLPGVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTYKMLYVSN 420  
Qy 548 KYLELAKLDLFDNMVQALHQLKQETQHIVSWWRSSGFNDLFTTRQRPVEMYFSVAVSMFEPEFA 607

Qy 781 GELVSSVQCYMRENPECTEESALSHVYGIIDNALKELNWEANPASNAPLCVRRLLFNTA 840  
Db 722 GELVSSVQCYMRENPECTEESALSHVYGIIDNALKELNWEANPASNAPLCVRRLLFNTA 781  
Qy 841 RMQLFYMYRDGFGISDKEMKHVSRTLPDPA 873  
Db 782 RMQLFYMYRDGFGISDKEMKHVSRTLPDPA 814  
RESULT 5  
US-10-041-007-37  
; Sequence 37, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 37  
; LENGTH: 795  
; TYPE: PRT  
; ORGANISM: Ginkgo biloba  
US-10-041-007-37  
Query Match 90.9%; Score 4224; DB 13; Length 795;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 80 LNADYHPAVWKDFIDSLTSPNSHATSKSVDTINKRIQTLVKEIQCMFQSGMDGETNP 139  
Db 2 LNADYHPAVWKDFIDSLTSPNSHATSKSVDTINKRIQTLVKEIQCMFQSGMDGETNP 61  
Qy 140 SAYDTAWARIPSIDGSGAPQPTQLQWLNQPLDGSWGSECFILAYDRVLNTLACLLT 199  
Db 62 SAYDTAWARIPSIDGSGAPQPTQLQWLNQPLDGSWGSECFILAYDRVLNTLACLLT 121  
Qy 200 LKTNKNGDIOVQKGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEAKSLGLDLPYHL 259  
Db 122 LKTNKNGDIOVQKGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEAKSLGLDLPYHL 181  
Qy 260 FISQIHQKQKLOKIPLVNLHNTALLYSLGLQDVVDWQEIITNLQSRDGSFLSSPAS 319  
Db 182 FISQIHQKQKLOKIPLVNLHNTALLYSLGLQDVVDWQEIITNLQSRDGSFLSSPAS 241  
Qy 320 TACVFMHTQNKRCCLHFLNVLKFGDYVPCHPYPLDLFRLWADTVRILGIDRYFKKEIK 379  
Db 242 TACVFMHTQNKRCCLHFLNVLKFGDYVPCHPYPLDLFRLWADTVRILGIDRYFKKEIK 301  
Qy 380 ESLDYVYRYWDAERGVCWACNPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEKGF 439  
Db 302 ESLDYVYRYWDAERGVCWACNPIPDVDDTAMGLRILRLHGYNVSSDVLENFRDEKGF 361  
Qy 440 CFAGOTQIGVTDNLNLYRCSQVCFPEKIMEEAKFTTTHLQNALAKNAPDKWAKKOL 499  
Db 362 CFAGOTQIGVTDNLNLYRCSQVCFPEKIMEEAKFTTTHLQNALAKNAPDKWAKKOL 421  
Qy 500 PGEVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTYKMLYVSNKYLELAKLDFNM 559  
Db 422 PGEVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTYKMLYVSNKYLELAKLDFNM 481  
Qy 560 VQALHQLKQETQHIVSWWRSSGFNDLFTTRQRPVEMYFSVAVSMFEPEFACTIAKTSCL 619  
Db 482 VQALHQLKQETQHIVSWWRSSGFNDLFTTRQRPVEMYFSVAVSMFEPEFACTIAKTSCL 541  
Qy 620 AVILDDLTDHSGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLYNTVNGFGKGLK 679



```
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/900,797.
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US/09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Abies grandis
; UN-09-900-797-56
```

```
Query Match      61.2%; Score 2844; DB 10; Length 868;
Best Local Similarity 62.3%; Pred. No. 5.9e-262;
Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;
```

```
QY   12 SLQLSPKVP---FRQSTNLIIPFHKRSGFNGNAQHCVSRHLRWLN-----CVGIIHA 60
    :|::||::||::||::||::||::||::||::||::||::||::||::||::||:
Db   5 SSSLSQQIPTAAHHLTANAQSIPHFTSLTLAGSSAKRRSLYLWKGSKNKIACV--E 62

QY   61 SAATRPDLPQEERFYVR---LNAVYPHAVWKDDFDISLTSPNSHATSXSVDETINKR 117
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   63 GGATSVPYQSAEKNDLSSTLVKRBPFGFWKKDDLIDSLTS--SHKVAAS--DE---KR 115

QY   118 IQTLVKETIQCFMQSGMGETNPASYDTAWARIPSIDGSGAPOPQTLOWILANNQLPDGS 177
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   116 IETLSEIKMFRCMGYGETNPASAYDTAWARIPAVDGSDPNHPPEIVTEWLQNQLKDGS 175

QY   178 WGEECIFLAYDRVLNTLACLITLKINWGDIQVQGVEFVRKMMEEMKDEADNRPSGFE 237
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   176 WGEGFVFAYLRILLATLACIITLWRGTGETQVQKIEFFTRTAGMKWEADSHRPSGFE 235

QY   238 VYFPAMLDEAKSLGLDLPHYLPFTSQIHQBOKKLQKIPLNLVHNHOTALLYSLLEGQDV 297
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   236 IVFPAMLEAKTLGLDLDPFLKQIIKEKAALKRIPTDVDLYALPTRLTYSLGLEQEI 295

QY   298 VDWOEITNLQSDGSLSSPASTACVFMHTQNKRLFNPLNVLSKFQGYVPCHYPLDLFE 357
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   296 VDWOXIMKLQSKDGSFLSSPASTAATAAFVMRTGNKCKLDFLNFVLKFKGHNVPCHYPLDLFE 355

QY   358 RLWAVDTVERLGIDRYFKKETESLDYYRVYWDABERGVGWARCNPIPDVDDTANGRLIR 417
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   356 RLWAVDTVERLGIDRHFKEEIKEALDYVYSHWD--ERGIGWARENPVPDIDDITANGRLIR 414

QY   418 LHGVYNVSSDLNENFRDEXDKDFCFAGOTQIGVTDMNLNYRSQCVCFPQEKIMEEAKTFTT 477
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   415 LHGVYNVSSDLTKTFRDENGEPFCFLGGTQRGTMDMLNVNCRSHVSFPGETTMEEAKLCTE 474

QY   478 NHIQNALAKNNADFKNWVKDLPGEVEVALKYPMHRSMPRLARSYIEBQFSNDVWLKGT 537
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   475 RYLRNALENVDADFKNWPKKNIRGEVEVALKPWHKSMPRLARSYIENYGDPDWLWLGKT 534

QY   538 YYQMULVYSNEKEYLELAKLDFNNVQAALHQKETQHIVSWWREGSFNDLTTFRORPVEMYFSV 597
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   535 YNMPPYISNEKEYLELAKLDFNKVQSIHQTELQDLRRWWKSSGFTDLNFRREVTEIIFYSP 594

QY   598 AVSMFPEFAACRIAIAKTSCLAVIDLLDYDHGSLDDLKLFSEAARRWDISVLDSVRDN 657
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   595 ASFIPEPFESKREVVYTKTSNFTVLDDLYDAHGSLDDLKLFTESVKRWDSLVDQM-PQ 653

QY   658 QLVKVFGLGLYNTVNGFGKDLKEQGRDVLGRLKRWEGILASYTEAEWSAAKVPTPNE 717
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   654 QMKICFVGFTNTNDIAKEGRERQGRDVLGYIQNWVKVQLEAYTKEAESWAKEYVPSENE 713

QY   718 YVENAKVISIALATVNLINSIFFTCPELLPYILOQVDLRSKFLHLVSLTGRLINDTKTYOAE 777
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   714 YIENASVISIALGITVVLISALTFTCEVLTDVLSKIDRESRFQLMGLTGLRLVNDTKTYOAE 773

QY   778 RNRGELVSSVQCYMNRENPECTEBEALSXYGIIDNALKELNWELANPASNAPLCVRRLLF 837
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   774 RGQGEVASAIQCYMKOHKPKISEEALQHVYSVNENAIBELREFVN--NKIPDIYKRLVF 831

QY   838 NTARMVOLFYMRDGFGIS-DKEMKHORSRTLDPVA 873
    :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db   832 ETARIMQLFYMQGDGLTLSDHMKEIKHVKNCLFPVA 868
```

```
RESULT 9
US-09-900-797-56
; Sequence 56, Application US/09900797
; Publication No. US20030087406A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20030087406A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
```





Db 236 IVFPAMLEAKTLGDLDPVDFLLKQIIEKREAKUKRIPTDVLVYALPTLLYSLEGLQEI 295  
 Qy 298 VDWQEIITLQSRDGSFLSSPASTACVFMHTQKRCILHFNFLVSKFGDYVPCHYPLDLFE 357  
 Db 296 VDWQIKMLQSKDGSFLSSPASTAAVFMRTGNKKCLDFLNFVLFKFGNHVPCHYPLDLFE 355  
 Qy 358 RLWAVDTVERLIGIDRYFKKIEKESLDYVRYVDAERGVCWACNPIPDVDDTAMGLRIIL 417  
 Db 356 RLWAVDTVERLIGIDRHFKKEIEKALDYVYSHWD-ERGIWARENPVPDIDDTAMGLRIIL 414  
 Qy 418 LHGYNVSSDVLNFRDEKDFCFAGQTOIGVTNMLNLYRCQVCPGPPGKIMEEAKTFTT 477  
 Db 415 LHGYNVSSDVLNFRDENGECFCFLGQTOIGVTNMLNLYRCQVCPGPPGKIMEEAKTCTE 474  
 Qy 478 NHLONALAKNAPDKWAVKODLPGEVEYAIKYPWHRSMPRLBARSYIEOFGSNDVWLKGT 537  
 Db 475 RYLNALENVDAFDKWAFFKNIRGEVEYALKYPWHKSMRPLEARSYIENYGPDVWLKGT 534  
 Qy 538 VYKMLVSNKYLELAKLDPNMQALHOKETOHIIVSMWRESGNDLTETRORPVEMYFSV 597  
 Db 535 VYMPYISNEKYLELAKLDPNMQALHOKETOHIIVSMWRESGNDLTETRORPVEMYFSV 594  
 Qy 598 AVSMPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDLDLKLFPSEAVRRWDISVLDSVRDN 657  
 Db 595 ASFIPEPEFSKCEVYTKTSNFTVILDDLYDAHGSLLDLDLKLFPSEAVRRWDISVLDSVRDN 653  
 Qy 658 QLVKVCFLGLYNTVNGFGKGLKEQGRDVLGYLKWEGGLASVYKAEWSAAKYVPTFNE 717  
 Db 654 QMKICFVGFYNTFNDIAKEGRERQGRDVLGYIQNVKWLQVLEAYTKEAWESEAKYVPSFNE 713  
 Qy 718 YVENAKVSTALATVNLNSIFFTGCELLPDYILOQVDRSKFLHVLSTLGRINDTKTYQAE 777  
 Db 714 YIENASVSIAGTVNLISALFTGEVLTDEVLSKIDRESRLQLMGLTGLVNDTKTYQAE 773  
 Qy 778 RNRGELVSSVQCYMRENPECTEAEALSHVYGIIDNALKEINWELANPASNAPLCVRRLIF 837  
 Db 774 RQGEVASAIQCYMKDHPKISEBEALQHVYSVMENALEELNREPVN--NKIPDIYKRLVF 831  
 Qy 838 NTARVMOLFMYNRDGFIS-DKEMKDHVSRTLDPVA 873  
 Db 832 ETARIMQLFYMQGDLTLSDHMEIKHVKNCCLFQVPA 868

RESULT 12  
 US-10-041-007-4  
 ; Sequence 4, Application US/10041007  
 ; Publication No. US20020164736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matsuda, Seiichi P.T.  
 ; APPLICANT: Schepmann, Hala G  
 ; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
 ; FILE REFERENCE: P02081US1  
 ; CURRENT APPLICATION NUMBER: US/10/041,007  
 ; PRIOR FILING DATE: 2002-01-07  
 ; PRIOR FILING DATE: 2001-01-05  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 868  
 ; TYPE: PRT  
 ; ORGANISM: Abies grandis  
 US-10-041-007-4

Query Match 61.2%; Score 2844; DB 13; Length 868;  
 Best Local Similarity 62.3%; Pred No. 5, 9e-262;  
 Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;  
 Qy 12 SLQLSPKVP---FROSTNLIPIFKRSGFGNAQCVSHLRLWN-----CVGIHA 60  
 Db 5 SSSLSSQIPTAHILTANAQSIPIPHFTSTLINAGSSAKRSRLRWGKSNKIACVGV--E 62

Qy 61 SAAETRPDQLPOEBRFVSR---LNADYHPAVWKDDFTDSLTPNSHATSKSVDETINKR 117  
 Db 63 GGATSVYQSAEKNDLSLSSSTLVKREFPPGFGKDDLDLSLTS--SHKVAAS--DE---KR 115  
 Qy 118 IOTLVKELTOCHQFOSNGOGETNPSAYDTAWARIIPSIDSGAPQPOFTQIQLWINNQLPDGS 177  
 Db 116 IETLISEIKNFRCMGYGETNPSAYDTAWARIIPAVDGSNDPHPEFTVEMILQNLKDG 175  
 Qy 178 WGEICIFLAYDRVLNLTACLTLKIWNKGDIOVQKGVFVRKHEEMKDEADNHRPSGFE 237  
 Db 176 WGEYFYFLAYDRILLATLACIITLWRTGETQVKGIEFFRTQAGKMEDEADSHRPSGFE 235  
 Qy 238 VYFPAMLEAKSLGDLDPYHLPFISQIHKQKQKLOKTPLVNLHNNHQTALLYSLGLQDV 297  
 Db 236 IVFPAMLEAKSLGDLDPYHLPFISQIHKQKQKLOKTPLVNLHNNHQTALLYSLGLQEI 295  
 Qy 298 VDWQEIITLQSRDGSFLSSPASTACVFMHTQKRCILHFNFLVSKFGDYVPCHYPLDLFE 357  
 Db 296 VDWQIKMLQSKDGSFLSSPASTAAVFMRTGNKKCLDFLNFVLFKFGNHVPCHYPLDLFE 355  
 Qy 358 RLWAVDTVERLIGIDRYFKKIEKESLDYVRYVDAERGVCWACNPIPDVDDTAMGLRIIL 417  
 Db 356 RLWAVDTVERLIGIDRHFKKEIEKALDYVYSHWD-ERGIWARENPVPDIDDTAMGLRIIL 414  
 Qy 418 LHGYNVSSDVLNFRDEKDFCFAGQTOIGVTNMLNLYRCQVCPGPPGKIMEEAKTFTT 477  
 Db 415 LHGYNVSSDVLNFRDENGECFCFLGQTOIGVTNMLNLYRCQVCPGPPGKIMEEAKTCTE 474  
 Qy 478 NHLONALAKNAPDKWAVKODLPGEVEYAIKYPWHRSMPRLBARSYIEOFGSNDVWLKGT 537  
 Db 475 RYLNALENVDAFDKWAFFKNIRGEVEYALKYPWHKSMRPLEARSYIENYGPDVWLKGT 534  
 Qy 538 VYKMLVSNKYLELAKLDPNMQALHOKETOHIIVSMWRESGNDLTETRORPVEMYFSV 597  
 Db 535 VYMPYISNEKYLELAKLDPNMQALHOKETOHIIVSMWRESGNDLTETRORPVEMYFSV 594  
 Qy 598 AVSMPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDLDLKLFPSEAVRRWDISVLDSVRDN 657  
 Db 595 ASFIPEPEFSKCEVYTKTSNFTVILDDLYDAHGSLLDLDLKLFPSEAVRRWDISVLDSVRDN 653  
 Qy 658 QLVKVCFLGLYNTVNGFGKGLKEQGRDVLGYLKWEGGLASVYKAEWSAAKYVPTFNE 717  
 Db 654 QMKICFVGFYNTFNDIAKEGRERQGRDVLGYIQNVKWLQVLEAYTKEAWESEAKYVPSFNE 713  
 Qy 718 YVENAKVSTALATVNLNSIFFTGCELLPDYILOQVDRSKFLHVLSTLGRINDTKTYQAE 777  
 Db 714 YIENASVSIAGTVNLISALFTGEVLTDEVLSKIDRESRLQLMGLTGLVNDTKTYQAE 773  
 Qy 778 RNRGELVSSVQCYMRENPECTEAEALSHVYGIIDNALKEINWELANPASNAPLCVRRLIF 837  
 Db 774 RQGEVASAIQCYMKDHPKISEBEALQHVYSVMENALEELNREPVN--NKIPDIYKRLVF 831  
 Qy 838 NTARVMOLFMYNRDGFIS-DKEMKDHVSRTLDPVA 873  
 Db 832 ETARIMQLFYMQGDLTLSDHMEIKHVKNCCLFQVPA 868

RESULT 13  
 US-10-041-007-15  
 ; Sequence 15, Application US/10041007  
 ; Publication No. US20020164736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matsuda, Seiichi P.T.  
 ; APPLICANT: Schepmann, Hala G  
 ; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
 ; FILE REFERENCE: P02081US1  
 ; CURRENT APPLICATION NUMBER: US/10/041,007  
 ; PRIOR FILING DATE: 2002-01-07  
 ; PRIOR FILING DATE: 2001-01-05  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 15

```
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Abies grandis
US-10-041-007-15

Query Match      45.5%; Score 2114.5; DB 13; Length 817;
Best Local Similarity 49.5%; Pred. No. 2.9e-192;
Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;

QY 79 RLNADHPAVKDDFIDSUTSPNSHATSKSSVDETINKRIQTLVKEIQCMFQS--MGDE 136
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 25 RRTANPHNVWGDLVHSLKSP-----YIDSSYRERAEVLVSEIKAMLNPAITGDGE 76

QY 137 T--NPSAYDTAWARIPSIDSGAPQPTQWLNNQLPDGSGEGECIFLAYDRVLNTL 194
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 77 SMITPSAYDTAWARIPSIDSGAPQPTQWLNNQLPDGSGEGECIFLAYDRVLNTL 136

QY 195 ACILTLKIWNKGDIQVQKGVFVRKHMEEMKDEAD--NHRPSGFVFPVFPAMLEAKSLGID 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 137 SCVVLVLLKWNVGDVQVEGIEIKSNLELVKDETDQDSLVTFEIIFFSLLREASLRIG 196

QY 254 LPVHLPFISQIHQKQKLOKILNVLNHNQHTALLYSLGLEQDVVDVQBEITLQSRDGSF 313
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 197 LPYDLPYIHLQTKQERLAKLSREIYAVPSFLYSLGIIQIVEMERIMEVQSDGSF 256

QY 314 LSSPASTACVFVHTQKRCGLHFLNPLVLSKFGDVVCHYPLDLPERLMAVDVTERIGIDRY 373
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 257 LSSPASTACVFVHTQKRCGLHFLNPLVLSKFGDVVCHYPLDLPERLMAVDVTERIGIDRY 316

QY 374 FKKEIKESLDYVRYVWDAERGVGWARCNPIPDVDDTAMGLRLRLHGVNVSVDLENFRD 433
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 317 FEKEIKEALDYVYRHN--ERGIWGLNPIADLETTALGFRLLRLHRYNVSIPAIFDNFKD 375

QY 434 EKGDFPCFAGQTOIGVTDNLNLYRCSQVCPGKEIMEEAKTFTTNHLQNALAKNNAFDM 493
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 376 ANGKFCSTGQFQKDVASMLNLYRASQLAPFGENILDEAKSPATKYREALSEKSTSSAW 435

QY 494 AVKKDLPGEVEVAKYPWHRSMPLREARSYIEQFGSNDVLMGKTVYKMLYVSNKYLELA 553
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 436 NNKQNSLBQIKYALKTSHASVPVREAKRYCQVYRDPYARIACVYKLPYVYNNKFLBEG 495

QY 554 KLDFFNMVQALHOKETOHIVSWWRESGNDLTFTRQRPVEMYSFVAVSMFPEFAACRIAY 613
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 496 KLDFFNIQSIHQEEMKNVTSWFRDGLPLFTFARERPLEFYLVAAGYEPQYAKCRFLP 555

QY 614 AKTSLAVLDDLDPYTHGSLDDLKLFSEAVRRWDISVLDSVRDNLQKVCFLGLYNTVNGF 673
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 556 TKVACLQTVLDDMDYDTYGTDLBELKFTAEVRRWDLSTENLPD--YMKLCYQIYYDIHVEV 614

QY 674 GKDGLKEQGRDVLGYLRKWEGILLASYTEAEWSAAKYVPTFNEVENAKVSIATVVL 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 615 AWEAEKEQGRDVLGYLRKWEGILLASYTEAEWSAAKYVPTFNEVENAKVSIATVVL 733

QY 734 NSI--PPTGELLPDYILQQVDL--RSKFLHLVSLTGLRLNDTKTYQAEARNRGELVSSVOCY 790
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 675 SGVLMDGQLLSQAELEKVDYVGRVLTSLNLSRLADDTKTYKAEKARGELASSIECY 734

QY 791 MRENPECTEERALSHYGIIDNALKELNWLANPASNAFLCVRLLFTNARVMQLFYMYR 850
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 735 MKDHPCTEERALSHYGIIDNALKELNWLANPASNAFLCVRLLFTNARVMQLFYMYR 850

QY 851 DGFGLSDKEMKDHVSRTELPDPV 872
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 794 DGFGLSVKLEVKDHIKECLTEPL 815
```

## RESULT 14

```
US-09-887-586A-46
; Sequence 46, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A11, Joseph P.
```

```
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-887-586A-46
```

```
Query Match      44.9%; Score 2085.5; DB 9; Length 782;
Best Local Similarity 49.6%; Pred. No. 1.6e-189;
Matches 391; Conservative 158; Mismatches 221; Indels 19; Gaps 9;

QY 92 DFIDSLTSPNSHATSKSSVDETINKRIQTLVKEIQCMFQS--MGDE 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 3 DLVHSLKSP-----YIDSSYRERAEVLVSEIKVMLNPAITGDGSMITPSAYDTAW 54

QY 148 ARIPSIDSGAPQPTQWLNNQLPDGSGEGECIFLAYDRVLNTLACILTLKIWNKGD 207
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 55 ARVPAIDGSARPQPTQWLNNQLPDGSGEGECIFLAYDRVLNTLACILTLKIWNKGD 114

QY 208 IQVQKGVFVRKHMEEMKDEAD--NHRPSGFVFPVFPAMLEAKSLGIDLPYHLPFISQIHQ 266
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 LQVEGIEIKSNLELVKDETDQDSLVTFEIIFFSLLREASLRIGLPLPYIHLQ 174

QY 267 KROKKLOKILNVLNHNQHTALLYSLGLEQDVVDVQBEITLQSRDGSFLSSPASTACVPMH 326
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 175 KROKKLOKILNVLNHNQHTALLYSLGLEQDVVDVQBEITLQSRDGSFLSSPASTACVPMH 234

QY 327 TONKRCRLHFLNVLKSGDVVCHYPLDLPERLMAVDVTERIGIDRYVKEIKESLDYVY 386
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 235 TGDACKLEFLNLMVKNFVPCLYPDVLLERLLIYVNLVGLIYRHEKIKESLDYVY 294

QY 387 RYDAERGVGWARCNPIPDVDDTAMGLRLRLHGVNVSVDLENFRDSEKGFPCFAGQ 446
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 295 RHWN--ERGIWGLNPIADLETTALGFRLLRLHRYNVSIPAIFDNFKDANGKFCSTGQFN 353

QY 447 IGVTDNLNYRCSQVCPGKEIMEEAKTFTTNHLQNALAKNNAFDMKAVKDLPGVEVYA 506
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 KDVASMLNLYRASQLAPFGENILDEAKSPATKYREALSEKSTSSAMNNKQNSLBQIKY 413

QY 507 IKYPWHRSMPLREARSYIEQFGSNDVLMGKTVYKMLYVSNKYLELAKLDFNMVQALHOK 566
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 414 LKTSWHLASVPVREAKRYCQVYRDPYARIACVYKLPYVYNNKFLBEGKLDFFNIQSIHQ 473

QY 567 ETQHIVSWWRESGNDLTFTRQRPVEMYSFVAVSMFPEFAACRIAYAKTSLAVLDDL 626
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 474 EMKNVTSWFRDGLPLFTFARERPLEFYLVAAGYEPQYAKCRFLPTKVAACLTQVLDDM 533

QY 627 YDTHGSLDDLKLFSEAVRRWDISVLDSVRDNLQKVCFLGLYNTVNGFGKGLKEQGRDVL 686
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 534 YDTYGTDLBELKFTAEVRRWDLSTENLPD--YMKLCYQIYYDIHVEVWEAEKSGRGLV 592

QY 687 GYLKRWEGILLASYTEAEWSAAKYVPTFNEVENAKVSIATVVLNLSI--FFTGELLPD 745
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 593 SFFRKGHEDYLLGYEEAEWLAAYVTLDEYIKNGITSIGORILLSGVLMDGQLLSQ 652

QY 746 YILQQVDL--RSKFLHLVSLTGLRLNDTKTYQAEARNRGELVSSVQCYMRENPECTEERAL 803
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 653 EALSVDYVGRVLTSLNLSRLADDTKTYKAEKARGELASSIECYMKDHPCTEERAL 712

QY 804 SHVGIIDNALKELNWLANPASNAFLCVRLLFTNARVMQLFYMYRDPGFSIKEMKDH 863
```

Db

713

DIHYSILEPAVKELTREFLKP-DDVPFACKKMLFETRTVMTWIFKDGDFGVSKLEVKDH

771

QY

864

VSRITLDPV

872

Db

772

IKECLIEPL

780

RESULT 15

US-09-903-012-46

; Sequence 46, Application US/09903012

; Patent No. US20020094557A1

; GENERAL INFORMATION:

; APPLICANT: Chappell, Joseph

; APPLICANT: No. US20020094557A1, Joseph P.

; APPLICANT: Starks, Courtney M.

; APPLICANT: Manna, Kathleen R.

; TITLE OF INVENTION: SYNTHASES

; FILE REFERENCE: 07678-025001

; CURRENT APPLICATION NUMBER: US/09/903,012

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 09/398,395

; PRIOR FILING DATE: 1999-09-17

; PRIOR APPLICATION NUMBER: 60/100,993

; PRIOR FILING DATE: 1998-09-18

; PRIOR APPLICATION NUMBER: 60/130,628

; PRIOR FILING DATE: 1999-04-22

; PRIOR APPLICATION NUMBER: 60/150,262

; PRIOR FILING DATE: 1999-08-23

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 46

; LENGTH: 782

; TYPE: PRT

; ORGANISM: Abies grandis

US-09-903-012-46

Query Match

44.9%; Score 2085.5; DB 9; Length 782;

Best Local Similarity

49.6%; Pred. No. 1.6e-189;

Matches 391; Conservative 158; Mismatches 221; Indels 19; Gaps 9;

QY

92

DFIDSLTSPNSHATSKSVDETINKRIQTLVKEIQCMFQS--MGDGET--NPSAYDTAWV

147

Db

3

DLVHSLKSP-----YIDSSYRERAEVLVSEIKWMLNPAITGDGSMITPSSAYDTAWV

54

QY

148

ARIPSIDGSGAPQPTQLQWLNQLPDGSWGECIFLAYDRLNTLACLTLKIWNKGD

207

Db

55

ARVPAIDGSGARPQPTQVDWILKNQPKDGSWGIQSHFLLSDRLATLSCVLVLLKNNVG

114

QY

208

IQVOKGVFVRKHEEMKDEAD-NHRPSGFVFPAPMLDEAKSLGLDLPYHLPIFISQIHQ

266

Db

115

LQVEQIEFIKSNLELVKDETDQDSLVTDFFIIFPSLLREAOQLRLGLPYDLPYHLLQT

174

QY

267

KRQKXLOKIPILNVLNHNHTALLYSLEGLQDVVDWQEIYNLQSRDGSFLSSPASTACVPMH

326

Db

175

KQOERLAKLSREBIYAVPSPLLYSLEGIQDIVEMERIMEVQSGDSFLSSPASTACVPMH

234

QY

327

TQNKCELFNLFVLSFGDYVPCYPLDLFERLWADVTVRGLIDIRYFKGKIKESLDYVY

386

Db

235

TGDACLEFLNSVMIKFNFVCLPVDLLERLLIVDNIIVRLGIYRHFKEIKEALDIVY

294

QY

387

RYWDAERGVMGARNCPIDVDVDTAMGLRILRLHGNVNSDVLNENRDEKGDFFCFAGQTO

446

Db

295

RHWN-ERGIGWGLNPIADLETTALGLRLLRLHRYNVSPAI PDNFKDANGKFCSTGQFN

353

QY

447

IGVTNINLYRSCVCFCEKIMEEAKTFTTNHLQNALAKNNAFDKWAVKDLPGVEVYA

506

Db

354

KOVASMLNLYRASQAFPGENILDEAKSPATKYLEALEKSETSSAANNKQNLQSEIKYA

413

QY

507

IKYPWHRSMRLEARSYIEQFGSNDVWLKGTYYKMLYVSNKEYLELAKLDFNMVQALHOK

566

Db

414

LKTSWHASVPRVEAKRYCQVYRPDYARIKCVVKLPYVNNKFLFELGKLDNFNIIQSIHQE

473

Search completed: August 17, 2004, 21:12:39

Job time : 50.4515 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:57:35 ; Search time 17.2551 Seconds  
(without alignments)  
4866.690 Million cell updates/sec

Title: US-10-041-007-2  
Perfect score: 4645  
Sequence: 1 MAGVLFANLPCSLQLSPKVP.....GISDKMKDHVSRTLDPDVA 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1332	28.7	802	2 D85035	ent-kaurene synthase
2	1318	28.4	823	2 T02959	kaurene synthase A
3	1303	28.1	801	2 T06783	ent-kaurene synthase
4	1072.5	23.1	789	2 T09672	ent-kaurene synthase
5	1034	22.3	785	2 T52059	ent-kaurene synthase
6	1033	22.2	785	2 G96825	hypothetical prote
7	705	15.2	870	2 A96637	vetispiradiene syn
8	663	14.3	520	2 A56118	limonene cyclase -
9	660	14.2	599	2 A48863	probable limonene
10	654.5	14.1	600	2 F71434	probable limonene
11	650.5	14.0	1024	2 G71434	(+)-delta-cadinene
12	647.5	13.9	554	2 S68365	(+)-delta-cadinene
13	643.5	13.9	554	2 S68366	5-epi-aristolochene
14	633.5	13.6	550	2 T03714	sequiterpene cycl
15	633.5	13.6	559	2 T08174	probable limonene
16	633	13.6	591	2 H84633	germacrene C synth
17	565	12.2	548	2 T06266	hypothetical prote
18	559	12.0	548	2 T06265	hypothetical prote
19	469.5	10.1	632	2 B36723	hypothetical prote
20	460.5	9.9	598	2 H86460	vetispiradiene syn
21	448	9.6	350	2 C56118	probable terpene s
22	443	9.5	608	2 G86443	hypothetical prote
23	416.5	9.0	573	2 T05328	hypothetical prote
24	403.5	8.7	582	2 C71424	hypothetical prote
25	393	8.5	530	2 G96588	hypothetical prote
26	389.5	8.4	535	2 T06285	hypothetical prote
27	398.5	8.4	421	2 C96642	hypothetical prote
28	375	8.1	598	2 T05329	probable vetispira
29	374	8.1	598	2 T00509	

30	371	8.0	501	2 H70972	probable cyclase -
31	369	7.9	598	2 F96684	probable terpene s
32	351.5	7.6	300	2 B56118	vetispiradiene syn
33	349.5	7.5	612	2 T05331	hypothetical prote
34	330.5	7.1	471	2 T06287	hypothetical prote
35	313	6.7	383	2 D71424	hypothetical prote
36	271	5.8	952	2 JC7227	ent-kaurene synthase
37	250	5.4	946	2 T00024	ent-kaurene synthase
38	227.5	4.9	516	2 T10874	y4kr protein - Rhi
39	203.5	4.4	516	2 I40214	hypothetical prote
40	166	3.6	203	2 H96525	probable terpene c
41	151.5	3.3	148	2 T03982	5-epi-aristolochene
42	146	3.1	979	2 E72236	clostriopain-relate
43	135.5	2.9	1131	2 T30951	hypothetical prote
44	135.5	2.9	3864	2 D87757	protein C44B4.1a [
45	132	2.8	2324	1 A29924	acetyl-CoA carboxy

## ALIGNMENTS

### RESULT 1

D85035

ent-kaurene synthetase A-like protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001

C;Accession: D85035

R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spri

Nature 402, 769-777, 1999

A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A;Reference number: A85001; MUID:20083488; PMID:10617198

A;Accession: D85035

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-802 <STO>

A;Cross-references: GB:NC\_001268; NID:g7269763; PIDN:CAB77763.1; GSPDB:GN00140

C;Genetics:

A;Gene: AY4902780

A;Map position: 4

Query Match 28.7%; Score 1332; DB 2; Length 802;  
Best Local Similarity 36.6%; Pred. No. 2e-83;  
Matches 283; Conservative 146; Mismatches 254; Indels 90; Gaps 17;

Qy	122	VKEIQCMFQSMGDETPNSAYDTAWVARIPSIDSGAPQPTQLQWILNNQLPDGSGEE	181
Db	98	VKSVKTLRNLDTGEITISAYDTAWVALIDA--GDKTPAPPSAVKWAENQLSDGSMGDA	155
Qy	182	CIFLAYDRLVLTACLLTLKWNKGDIOVKGVFVRKHEEMKDEADNHRPSPGVFVFP	241
Db	156	YLFSDYHRLTLTACVVALRSWNLFPHQCNKGITFFRENIGKLEDENDHMPIGFEVAFP	215
Qy	242	AMLDEAKSLGLDLPYHLPFISQIHQKROKLOKIPLNVLNHTALLYSLEGQDVVDQ	301
Db	216	SLLEIARGINIDVPYSPVLKDIYAKKELKLTIPKEIMKIPITLLHSLSEGHRD-LDWE	274
Qy	302	EITNLQRDGSFLSSPASTACVPMHTQNKRCFLNVLNFKDGYVPCHYPLDLFERLWA	361
Db	275	KLLKLQSDGSFLSPSSATAFMQTRDSNCLYLRNAVCRFNGVPNVFVDFLFEHIWI	334
Qy	362	VDTVRLGIDRYPKKEIKESLDYVYVWDARGVGWARCNIPIDVDTDATAGLILRHGY	421
Db	335	VDRLQRLGISRYFEEBIKECLDYVHYRW-TDNGICWARCASHVDIDDTANAFRLLRQHY	393
Qy	422	NVSSDVLNFRDSKGFPCFAGOTOIGVTDNLNLRYCSQVCFPEKIMEAKTFTTNHLQ	481
Db	394	QVSADVFKNF-EKEGEFFCFVGQSNQAVTGMFNLYRASQAFPRFELKNKGEFSYNLL	452
Qy	482	NALAKNAFDKVAWKDLDPGEVEYAIKYPWHRSMPRLARSYIEQF-GSNDVWLKGTYYK	540
Db	453	EKREEREELDKWIIMKDLPGEIGFALEIPWYASLPVETRFYIDQYGGENDVWIGKLYR	512
Qy	541	MLYVSNEKYLELAKLDFNMVQALHQKETOHIVSWRESGNLDTFTTRQRPVEMYFSVAVS	600

Db 513 MPYVNNNGYLELAKQDYNQAOHQLEWDFQKWEENRLSEWGRVRSSELCEVYLAAT 572  
Qy 601 MPEPEFAACRIAYAKTSCGLAVILDLDYTHGSLDLDLKFSEA-----VRRWDISVLD- 652  
Db 573 IFESERHERVMWAKSVKAISSFCG--SSDRSRFSDFQFHEYANARRSDHFNDR 630  
Qy 653 -----SVRDNOLKVCFLGLYNTVNGFGKDLKEQGRDVLGYLRLKVGWGLLASYTKEA 704  
Db 631 NMLRDRPGSVQASRLAGVLIG---TLNQMSFDLFWSHGRDWNLLYLSW-----G 677  
Qy 705 EWSAAKYVPTFENEYENAKVSIALATVYL---NSI--FFTCELLPDVILQOVDLRSKFLH 759  
Db 678 DM-----MEKWKLYGDEGEGLWKMIIILMNNDLTNFFT-----HTHFVR 718  
Qy 760 LVSLTGRINDTKTYQAEERNGELVSSVQCYMRENPECTEALSHVYGIIDNALKEINW 819  
Db 719 LAEINRICLPQYILKARN-----DEKTIKSMKEMKGMVELALSSED- 764  
Qy 820 ELANPASNAPLCVRRLLFNTARVMOLFMYRDFGFGISDKEMKHVSRFTLPDPV 872  
Db 765 -----TFRDVSITFLDVAKAFY-----FALCGDHLQTHISKVLFOKV 802

RESULT 2  
T02959  
kaurene synthase A - maize  
C:Species: Zea mays (maize)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 29-Oct-1999  
R:Accession: T02959  
R:Beneseon, R.J.; Johal, G.S.; Crane, V.C.; Tosberg, J.T.; Schnable, P.S.; Meeley, R.B.;  
Plant Cell 7, 75-84, 1995  
A:Title: Cloning and characterization of the maize An1 gene.  
A:Reference number: Z14794; MUID:95210929; PMID:7696880  
A:Accession: T02959  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-823 <BEN>  
A:Cross-references: EMBL:L37750; NID:g576885; PIDN:AAA73960.1; PID:g576886  
C:Genetics:  
A:Gene: AN1  
C:Function:  
A:Description: cyclase involved in gibberellic acid biosynthesis

Query Match 28.4%; Score 1318; DB 2; Length 823;  
Best Local Similarity 35.0%; Pred. No. 1.9e-82;  
Matches 302; Conservative 157; Mismatches 275; Indels 130; Gaps 26;  
Qy 71 PQERFVSRNLADYHPAVWKDDFDISLTPNSHATSKSSVDETINK----- 116  
Db 28 PQARRVVERAAGPGHATTQQ--PDNVSSAKVFTSRVETESKLRNGRKPODLEDEHQA 86  
Qy 117 --RIQTLVKEIQCMQSGMDGNPSAYDTAWARIPSI---DGSAPQPPQTLQWILNQ 172  
Db 87 EAELOPLIDQVAMRLSMNDGTSASAYDTAWAVWPKVGGDGAQPPFATVRIWIDHQ 146  
Qy 173 LPDGSGBECIFLAYDRVLTACLLTKIWNKGDIOVKGVFVRKHMEEMKDEADNR 232  
Db 147 LPDGSGBSALFSAIDRMINLTACVVALTKWSLPARCEAGLSFLHNMWRLAEERESM 206  
Qy 233 PSGEVVFVFPAMLDKASIG-LDLPHYLPFTSIQHKQKQLKQIPLNVLNHTALYSL 291  
Db 207 PIGEIAFPSSIQTARDLVVDVFFYPGHPALQSIYANREVKLRIPRDMHRVPTSLHSL 266  
Qy 292 EGLQDVMDQBITNQLQSDGSFLSPASTACVFMHTQNKRLHFLNVLKFGDVPVCHY 351  
Db 267 EGMPD-LDWPRLLNLQSCDGSFLSPSATAVALMQTGDKKCFEYIDRVKFGGVENVY 325  
Qy 352 PLDLFERLWAVDTVERLIGIDRYFKKEIKESLDIYVYRWDAERGVGWACNPIPDVDDTAM 411  
Db 326 PVDLFEHLVWDRLERLGISRYFQREIEQCMDYVNRHW--TEDGICWARKSNVQKVDVDTAM 384  
Qy 412 GLRILRLHGYNVSSVLENFRDEKGDFFCFAGQIQIGVTDNLNLYRCQSVQCFPGKIMEE 471

Db 385 AFLLRLHGYNVSPSPKFN-EKDGFFCFVGOSTQAVTGMYNLNRASQISFGEDVLRH 443  
Qy 472 AKTFTTNHONALAKONADFKWAVKDLGCEVEYAIKYPWHRSMRPLEARSYIEQF-GSN 530  
Db 444 ARVSEYFLRQREBEQMDIKWIVAKDLFGEVQYITLDFPWYASLPFRVEARTYLDQYGGKD 503  
Qy 531 DWMLGKTVYKMLVSVNEKYLELAKLDFNMVQALHQKETOHIVSWMRSENGNDLTFTRQRP 590  
Db 504 DWIGIKTLVYRPLVNNNDYLELAIRDFNHCOALHQECNGLQTYKDKNCLDAFGVEPQDV 563  
Qy 591 VEMYFVSVMFPEFAACRIAYAKTSCGLAVILDLDYTHGSLDLDLKFSEAARRWDISV 650  
Db 564 LRSYFLAAACIFEPSRAAERLAWARTSMIA---NAISTH-----LRDISDKRLSCFV 614  
Qy 651 LDSVRDNLK-----VCFGLYNTVNGFGKGL--KEQGRDVLGYLRLKVGWGLLAS 699  
Db 615 HCLYEENDSVLKRNPNDVILERLALRLINLLAQEALPIHEGQRFIHSLLSLAW----- 668  
Qy 700 YTKAEASAAKYVPTFENEYENAKVSIALATVVLNSIFFTGELLPDVILQOVDLRSKFLH 759  
Db 669 ----TEWMLQKANKENKHKSCGIE-----PQTMVHD---RQTYLL 703  
Qy 760 LVSL-----TGRINDTKTYQAEERNGELVSSVQCYMRENPE-----ECTEERALSHVYGI 810  
Db 704 LVQVIEICAGRI-----GEAVSMIN--NKDNDWFIQITCATCDLSLNRMLLS 748  
Qy 811 DNALK--ELNW-----ELA-----NPASNAPLCVRRLLFNTARVWOLFMY 849  
Db 749 QDTWKEARINWIEKEITELNQELAQSLLRCDKTSNKK--TKKTLMDVLR--SLYYAT 804  
Qy 850 RDGFGISDKEMKD-HVSRFTLPDPV 872  
Db 805 H-----SPQHMDIRHVSRIPEPV 823

RESULT 3  
T06783  
ent-kaurene synthase A (EC 2.5.1.-) - garden pea  
C:Species: Pisum sativum (garden pea)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: T06783  
R:Ait-Ali, T.; Swain, S.M.; Reid, J.B.; Sun, T.P.; Kamiya, Y.  
Plant J. 11, 443-454, 1997  
A:Title: The LS locus of pea encodes the gibberellin biosynthesis enzyme ent-kaurene sy-  
A:Reference number: Z15810; MUID:97260358; PMID:9107034  
A:Accession: T06783  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-801 <AIT>  
A:Cross-references: EMBL:U63652; NID:g2160543; PIDN:AAB58822.1; PID:g2160544  
A:Experimental source: cultivar Torsdag  
C:Genetics:  
A:Gene: LS  
C:Function:  
A:Description: converts the geranylgeranyl diphosphate into copalyl diphosphate  
A:Pathway: the first step of the gibberellin biosynthesis  
C:Keywords: transferase

Query Match 28.1%; Score 1303; DB 2; Length 801;  
Best Local Similarity 34.2%; Pred. No. 1.9e-81;  
Matches 297; Conservative 171; Mismatches 307; Indels 94; Gaps 22;  
Qy 19 VPRFQSTNLIIPHKSSSGFNQAHCVRSHLRWNCVGIHAGAAETRPDQLPQEEFVS 78  
Db 11 LPSSSSLFFLHPYKSSSLG-----AVSFVAKDKKRC-----RAIS 47  
Qy 79 RLNADYHPAYWKDDFDISLTPNSHATSKSSVDETINKRIQT-LYKEIQCMFQSGMGET 137  
Db 48 KSRTQEGEVFQTN-VATLKLSEINVEDVIVIDDEEQDIRVGLVKNKISLSLEDEGI 106  
Qy 138 NPSAYDTAWARIPSIDSGAGAPQPPQTLQWILNQLPDGSGWGECEIFLAYDRVLTACL 197

Db 107 TISAYDTAWALVEDVNAISTPQPPSLEWIAKQLODQSGWDSKRLPSAHDRIINTLACV 166  
 Qy 198 LTLKTNKGDIOVQKGVFVRKHEEMKDEADNRHPSQFVFPVPAAMLDEAKSL-GLDLPY 256  
 Db 167 IALRSWNNHSEKCDKMPFFRENLSKLENEHEMPIGFVAPFSLLEGARGIKPLMCPN 226  
 Qy 257 HLPFISQIHQKROKLOKIPLNVLHNOTALLYSLEGLODVVDQEIINLQSRDGSFLSS 316  
 Db 227 DSPILKNIPKREDEKLTIPKXIMHKVPTTLHLSLEGMSG-LDWKQLKQSQSGSFLPS 285  
 Qy 317 PASTACVPMHTQNRKCLHFLNVLKFGDYVYPCHYPLDPLPERLWAVDTVERLIGDRYFKK 376  
 Db 286 PSSTAFALMOTKQGNCLKYLANNVVKFNGGVVYVVDLPEHIVWVDRLERLIGSRFRH 345  
 Qy 377 EIKESLDVYRYWDAERGVGWACNPIPDVDDTAMGLRILHNGVSSDVLENFRDEKG 436  
 Db 346 EIKCMNVYSKIW-SEKIGICWARSNVQIDIDTAMAFELRLHGHVSAHFKEH-ERNG 403  
 Qy 437 DFFCFAGOTQIGVTDNLNLYRCSCVCPGKEIMEEAKTFTTNHNLQNALAKNAPDKWAVK 496  
 Db 404 EPPCFAGOTQAVTGMVYLFASQVLPFGKILHAKHFSKVLKREKANEELDKWIM 463  
 Qy 497 KDLPGVEYAIKYPMHRSMPLEARSYLEQFGS-NDVWLGKTVYKMLYVSNKYLELAKL 555  
 Db 464 KNLPEEVGYALDMPWYANLDRITRFYIDQGAESDVWIGKTYRMAYVNNNNYLELAKL 523  
 Qy 556 DFNWVQALHOKETOHIIYSWRESGFNDLTFTRQRPVEMFVAVSMPEPEPAACRIAYAK 615  
 Db 524 DYNMCOAHLIEWNIQVWYLESRLGEBGLSKLALLAYFLATGISIPEPERSHERLAWK 583  
 Qy 616 TSLAVIL-----DDLVDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFL-GLYN 668  
 Db 584 TTALETIKVVRNEDL-----RKDFAKFNDHIDVRDYSIARWKNKTHHELVESLFA 638  
 Qy 669 TVNGFGKDKLQEQGRDVLGY-LRKVWEGLLASYTEAEWSAAKTVPTFNVEYENAKVISA 727  
 Db 639 TIGEISMDVRLSYGHEI-GYDMHCQKWLSSQSGDKCRGE-----AELLIQ 686  
 Qy 728 LATVVLNSIFFTGELLPDYILQVDLRSKFLHLVSLTGRINDTKTYQAEARNRGEVSSV 787  
 Db 687 IINLCN-----HWISEGSPMSQSTIQHLLQLTNSICHKLSQCYQKKE-----LKG 732  
 Qy 788 QCYMRENPCTEEBALSHVYGIIDNALKEWELANWELANWELANWELANWELANWELAN 847  
 Db 733 SC--QEN--ITNSEVESKQSLVQVFOK-----CPNDIDFN---VKNTFF 771  
 Qy 848 MYRDFGIS-----DKEMKDHVSRTLPDPV 872  
 Db 772 TIAKSFYAAFCDSRTINFHIAKVLFEKV 800

RESULT 4

T09672  
 ent-kaurene synthase B (EC 2.5.1.-) - winter squash  
 C:Species: Cucurbita maxima (winter squash)  
 C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
 C:Accession: T09672  
 R:Yamaguchi, S.; Saito, T.; Abe, H.; Yamane, H.; Murofushi, N.; Kamiya, Y.  
 A:Title: Molecular cloning and characterization of a cDNA encoding the gibberellin biosyn-  
 plant J. 10, 203-213, 1996  
 A:Reference number: 216814; MUID:96367664; PMID:8771778  
 A:Accession: T09672  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-789 <YAM>  
 A:Cross-references: EMBL:U43904; NID:g1431869; PIDN:AAB39482.1; PID:g1431870  
 A:Experimental source: immature seeds  
 C:Function:  
 A:Description: catalyzes the conversion of copalyl diphosphate to ent-kaurene  
 A:Pathway: gibberellin biosynthesis  
 A:Note: terpene cyclase  
 C:Keywords: transferase

Query Match 23.1%; Score 1072.5; DB 2; Length 789;  
 Best Local Similarity 32.1%; Pred. No. 1.3e-55;  
 Matches 254; Conservative 164; Mismatches 324; Indels 49; Gaps 19;  
 Qy 75 RFVSRNLNADYHPAVKDDDFDSDLSITSPNSHATSSKSSVDETINKRIQTLVKEBIQCFQSMGD 134  
 Db 11 RFASSSSSSASLIFPGVDVDTTK-----TGALHPEET-KERIKKLFDV----- 55  
 Qy 135 GETNPAGYDTAWAVRIPSIDGSGAPQPPQFLQWILNNQLPDGSGW-----BECIFLAYDRV 190  
 Db 56 -ELSVSAYDTAWAVRIPSPNSLNQPLFPPECINVLDSQHADGWSGLLHNDQLLKA--NL 112  
 Qy 191 LNTLACLLTLKINWKGDIQVQKGVFVRKHEEMKDEADNRHPSGPEVFPVPAAMLDEAKSL 250  
 Db 113 LSTLACVLTLLKRNIGHDMHMSKALDFTKSNIASATDE-NORSPVGFDTIIPFGIMEYAKDL 171  
 Qy 251 GLDPLPHLPFISQIHQKROKLOKIPLNVLHNOTALLYSLEGLODVVDQEIINLQSRD 310  
 Db 172 NLNLPLAPTNDVALVRKKELELSCRSN-SEGKAVLAYVSEGIKGLQDMVMVQYQRKN 230  
 Qy 311 GSFLSPASTACVPMHTQNRKCLHFLNVLKFGDYVYPCHYPLDPLPERLWAVDTVERLIG 370  
 Db 231 GSLFNSPSTTAAAFMRNDDGCFDYLRLSLKQFDSVPTTYPLDIARLHMDVSLQKFGI 290  
 Qy 371 DRYFKKEIKESLDVYRYW-DAERGVGWACNPIPDVDDTAMGLRILHNGVSSDVLEN 429  
 Db 291 ARHFKBEIRSVLDETTRCMMQSGE-----NIFLDASTCAMAFRLURVEGVDVSSDQLT 343  
 Qy 430 NFRDEKGDFF--CFAGOTQ-IGVTDNLNLYRCSCV-VCFPGEKIMEEAKTFTTNHNLQALA 485  
 Db 344 QFSB---DIFPNCGLGGVLDKDFGA--SLELYKASQIITHPDESLENINSWTSRFLKHGLS 398  
 Qy 486 KNAFQKAWVKD--LPGEVEYAIKYPMHRSMPLEARSYLEQFGSNDVWLGKTVYKMLY 543  
 Db 399 SDSV---WSDRTDSVVKQEAVALFPYNATLERLISKRAVESYSGDIVRISKSPYACLN 455  
 Qy 544 VSNKYLELAKLDPNMQALHOKETOHIIYSWRESGFNDLTFTRQRPVEMFVAVSMPE 603  
 Db 456 FGHODEFLAVEFDNTLQRIHLKELBELQWVVENKLDLKFRLHLGYCYFAAAATLTD 515  
 Qy 604 PEPAACRIAYAKTSLAVILDDLDDYTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFL 663  
 Db 516 PELHDAIWAQNGVLTITVDDFDYDGGSEEDLNLELVEKWDPDGEVGYCKDVEIVP 575  
 Qy 664 LGIYNTVNGKDKLQEQGRDVLGYLRKVWEGLLASYTEAEWSAAKTVPTFNVEYENAK 723  
 Db 576 LALHSTVCEIGRRALVWQGRSVNRNVDGWLALLKVMRKEAEWSNKNVSPMSGEYMEQAH 635  
 Qy 724 VSIATATVVLNSIFFTGELLPDYILQVDLRSKFLHLVSLTGRINDTKTYQAEARNRGE 783  
 Db 636 VSPALGPILLPLMLFFVGPKLSEEMIGSCYQ-KLYKLMSTAGRLKNDIRSYDRECKEGKL 694  
 Qy 784 VSSVQCYMRE-NPECTEEBALSHVYGIIDNALKEWELANWELANWELANWELANWELAN 842  
 Db 695 -NLSLWMDGGGVNVTKEEAIEAKGDFEIRAILGLVLQENTTIPRACKDLFWKLMSI 753  
 Qy 843 MQLFYMYRDFG 853  
 Db 754 VNLFYMEDDGY 764

RESULT 5

T52059  
 ent-kaurene synthase (EC 2.5.1.-) [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
 C:Accession: T52059  
 R:Yamaguchi, S.; Sun, T.; Kawade, H.; Kamiya, Y.  
 A:Title: The GA2 locus of Arabidopsis thaliana encodes ent-kaurene synthase of gibberel-  
 Plant Physiol. 116, 1271-1278, 1998  
 A:Reference number: 225926; MUID:98205064; PMID:9536043  
 A:Accession: T52059  
 A:Status: preliminary; translated from GB/EMBL/DBJ



C;Accession: G96825  
R;Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K. ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: G96825  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-785 <STO>  
A;Cross-references: GB:AE005173; MID:g4835764; PIDN:AAD30231.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: T8K14.12  
A;Map position: 1

Query Match 22.3%; Score 1034; DB 2; Length 785;  
Best Local Similarity 32.4%; Pred. No. 5.9e-63;  
Matches 259; Conservative 156; Mismatches 329; Indels 56; Gaps 20;

QY 98 TSPNSHATSKSSVDETINKR-----IQTLVKEIQCMFQSMGGETNPSAYDTAWARIPS 152  
DB 11 SSPIS-ATLERGLDSEVQTRANNVSFEQTEKIRKMLEKV---ELSVSAYDTSWAMVPS 66

QY 153 IDGGAPOFPQTLQWLNQLPDGSWG---PECIFLAYDRVLNTLACLLTKIWNKGDIQ 209  
DB 67 PSSQAPLPFQCVKWLNDQHEDGSWGLNDHDSKVDLSSTLASILALKWIGIGERQ 126

QY 210 VQGVFEVRKMEEMKDEADNHRPSGFVFPFAMLDKSLGLDLPYHLPIFSIQHQRQ 269  
DB 127 INKGQFIETLSALVTDET-IQKPTGFDIIFPGMIKYARDLNLTIPLGSEVVDDMIRKRD 185

QY 270 KKLQKIPNLVNLHNTALLYSLEGLOVDVWQEIINLQSRDGSFLSSPASTACVPMHTQN 329  
DB 186 LDLCDSSEKFSKGRAYLAYVLEGRNLKDWDLIVKYORKNGSLFDSPTATAAFTQGN 245

QY 330 KCLHFLNVLKFGDYVPCHYPLDLFERLMAVDTVVERIGIDRYFKGEIKESLDVYRYW 389  
DB 246 DGLRYLCSLLQKFEAAVPSVYFPQYARLSIIVTLESIGIDRDFTKTEIKSILDTYRYW 305

QY 390 ---DAERGVMGARNPDPVDDTAMGLRILRLHGNVSDVLENFRDEKGFECFAGTQ 446  
DB 306 LRGDEEICLDLTC-----ALAFRLLLAHGVDVSDPLKPFABESG-----FSDTL 351

QY 447 IGWTDN-----LNIYRCSQVCFPEKIMEEAKFTTNHLQNALA---KNNAFDKWAKKDL 499  
DB 352 EGYVKNVTFVLELFAKQ--SYPHESALKKQCCWKQYLEWELSSWVKTSDVKYLKK---407

QY 500 PGVEVAIKYPMHRSMPRLARSYIEQFGS--NDVWLKTVYKMLVYSNEKYLELAKLDF 557  
DB 408 --EVEDALAPPSYASLERSDHRKILN-GSAVENTRVTKTSYRLHNICTSDILKLAVDVF 464

QY 558 NMVQALHQKETQHVIVSWWRSGENDLTFTRQPVEMVESVAVSMFPEPAACRIAYAKTS 617  
DB 465 NQCQSIHREMERLDRWIVENRLQELKFAQKLAICYFSGAATLFPSELSDARISWAKG 524

QY 618 CLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNQKLCVFLGLYNTVNGFGKDG 677  
DB 525 VLTIVDDDFDVGSGKEELENLHLVEKDLNGVPEYSSSEHVEIIFSVLRDTILETGDKA 584

QY 678 LKEQGRDVLGVLKRWEGGLASYTKAEWSAAKVPTFNVEVENAKVSIATATVVLNSIF 737  
DB 585 FTYQGRNVTHIIVKIVLWDLKSLMREASWSDKSTPSLEDYMENAYISFALGPVILPATY 644

QY 738 FTGELLDPYILQVLDLR--SKFLHLSITGLINDTKTYQARNRGLSVSSVQCYMR-EN 794  
DB 645 LIGPPLPE---KTVDSHQYNQLYKLVSTMGRLLNDIQGFKRESAEGKL-NAVSLHMKHER 700

QY 795 PECTEEELSHVYGIIDNALKELN-WELANPASNAPLCVRRLLNTARVWOLFMYVDGF 853  
DB 701 DNRSEKVIIESMKGLAEKREELHKLVEEKGSVVPRECKEAFKNGSKVLNLFYRKDDGF 760

QY 854 GISDKEMKDHVSRTLPDVA 873  
DB 761 --TSNDLSLVKSVIYEPVS 778

RESULT 6  
G96825  
Hypothetical protein T8K14.12 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001





probable limonene cyclase - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: Columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 05-Dec-1998  
C:Accession: G71434  
R;Bent, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks, P.; Wedler, H.; Wedler, E.; Wambut, R.; Weitzenecker, T.; Pohl, T.M.; Tarryn, N.; Gielavanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B. Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech, erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Anand, C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: G71434  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1024 <BEV>  
A:Cross-references: GB:297341; NID:g2244991; PID:e327011; PID:g2245029  
C:Genetics:  
A:Map position: 4COP9-4G3845

Query Match 14.0%; Score 650.5; DB 2; Length 1024;  
Best Local Similarity 32.8%; Pred. No. 2.1e-36;  
Matches 179; Conservative 97; Mismatches 178; Indels 91; Gaps 21;  
Qy 355 LFERLMAVDTVERLIGDRYFKKEIKESLDYVY-----RYWDAERGVGWACNPIPDV 406  
Db 27 LLEQLLELITLQRLGVSHFQEIKETLTITNVHVKNVRAHKNRIDNR-----WG-----DL 77  
Qy 407 DDTAMGLRIILRHGY-----NVSSDVLENFRDRKGGFFCFAGQTQIGVTDNLNLYRCSQVCP 463  
Db 78 YATALEFRLLRQHDVDFGDNIGVDL--DOKDITG-----ILSLYEASLYST 120  
Qy 464 PGKIMEBAKFTTTHNLQNALAKN-NAPDKWAKVKKDLPGVEYAIKYPWHRSMPLREARS 522  
Db 121 RIDTKLESIIYTTKRLKPFVKNKETSILRR---WVHALEMPYHRRVGRLEARW 176  
Qy 523 YIEQFGNDVWLGTVTYKGLVYSNEKYLELAKLDFNMVQALHOKETQHIIVSWWRSGF-N 581  
Db 177 YIEVYGER-----HDMNPILLELAKLDFNFVQATHDELKSLSSWWSKTGLTK 224  
Qy 582 DLTTRORPVMYFSVAVSMPEPEFAACRIAYAKTSCLAVIDLDLYTHGSLDLDLKLFSB 641  
Db 225 HLDVFRDITGEGYSSVGVMYEPFAYHROMLTKVFMILTITDIDYIGTLESLQFTT 284  
Qy 642 AVREWDISVLDSVRDNLKVCFLGLYNTVNGFGDKLKEQGRDVLGYLRKRWEGLLASYT 701  
Db 285 IVEKWDVNRLEEL-PNYMKLCFLCLVNEINQIGYFVLDRKGFNIPVYKESWADMCTFL 343  
Qy 702 KEAWSAAKYVPTNVEYENAKVSIATVVLNSIIFTGELLPPYILQQVDLRSKFLHLV 761  
Db 344 KEAKWYKSGYKPNFEYMQNGWISSVPTILLH-LFC---LLSD---QTLDLILGSYNHVS 396  
Qy 762 SLTG-----RLNDR-----KTYQAEER-NRCELVSVCYMRNPECTEEALSHVY 807  
Db 397 VRSSATILRLANDLATSVSHTGFTYNTTEELARGDTMKSVQCHMHET-GASEAESRAYIQ 455  
Qy 808 GIIDNALKELNWLANPASNAPLCVRRLLF-----NTARVWQLPFMYRDRDGFGISDK-EMK 861  
Db 456 GIIGVAVMDLNLNWKSS-----CRLHQGFLEAANLGRVAQCYQYQVGDGHGCPDKAKTV 508  
Qy 862 DHVSR 866  
Db 509 NHVRR 513

RESULT 12  
S68365  
(+)-delta-cadinene synthase isozyme XC1 - Gossypium arboreum  
C:Species: Gossypium arboreum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: S68365

R;Chen, X.Y.; Chen, Y.; Heinstein, P.; Davison, V.J. Arch. Biochem. Biophys. 324, 255-266, 1995  
A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a cDNA from Gossypium arboreum  
A:Reference number: S68365; MUID:96132653; PMID:8554317  
A:Accession: S68365  
A:Molecule type: mRNA  
A:Residues: 1-554 <CHE>  
A:Cross-references: EMBL:U23205; NID:g1045311; PIDN:AAA93064.1; PID:g1045312  
A:Experimental source: cultivar Nanking  
C:Superfamily: vetispiradiene synthase 1  
C:Keywords: phytoalexin biosynthesis

Query Match 13.9%; Score 647.5; DB 2; Length 554;  
Best Local Similarity 30.9%; Pred. No. 1.3e-36;  
Matches 162; Conservative 103; Mismatches 210; Indels 49; Gaps 12;  
Qy 357 ERLWAVDVTVERLIGDRYFKKEIKESLDYVY-RYWDAERGVGWACNPIPDVDDTAMGLRI 415  
Db 72 OKLAFIDSVQRLGVSHYFTKEIEDELENIYHNNDAEN-----DLYTTSIRPRL 120  
Qy 416 LRLHGYNVSDVLENFRDEKDPFCFAGQTQIGVTDN-----LNLYRCSQVCFPEEKIMEE 471  
Db 121 LREHGYNVSCDVFNKPKFQEGNF-----KSSVTSVDRGLLELYQASVLRVHGEDILDE 173  
Qy 472 AKTFTTNHLQNALAKNAPDKWAKVKKDLPGVEYAIKYPWHRSMPLREARSYIEQFGSND 531  
Db 174 AISFTTHHLSLAVA-----SLDHPLSSEVSHALKQSIRRGLPVREARHYL----- 218  
Qy 532 VWLGKTVYKMLVYSNEKYLELAKLDFNMVQALHOKETQHIIVSWWRSGF-NDLTFTTRQP 590  
Db 219 -----SVYQDIESHNKALLEFAKIDFNMLQFLHRKSELSECRWKKOLDFOKLPFYARDV 273  
Qy 591 VEMYFSVAVSMPEPEFAACRIAYAKTSCLAVIDLDLYTHGSLDLDLKLFSVARRWDISV 650  
Db 274 VEGYFWISGVYFEPQYSLGRKMLTKVIMASIVDDTVDSVATVEELIPTYNAIERWDIKC 333  
Qy 651 LDSVRDNLKVCFLGLYNTVNGFGDKLKEQGRD-VLGYLRKRWEGLLASYTEAEWSAA 709  
Db 334 IDEIBE-YMKPSYKALLDVYEM-VOLVAEHGQRYEYAKNAMIRLAQSYLYEAKWTQ 391  
Qy 710 KYVPTNVEYENAKVSIATVVLNSIIFTGELLPPYILQQVDLRSKFLHLVSLTGLIN 769  
Db 392 NYKPSPEEFKANALPTCGYAMLAITSFVNGMDIVTETPKWAASDPKIIQASIIICPFMD 451  
Qy 770 DTKTYQAEERNRGLSVSVOCYMRNPECTEEALSHVYIGIIDNALKELNWLANPASNAP 829  
Db 452 DVAEHKFKRRRDCSAIECYMEY-GVTAQAEYDVFNKHVESAWKDLNQEFLLKP-TEMP 509  
Qy 830 LCVRRLLFNFTARVWQLPFMYRDRDGFGISDKEMKHVSRTLPDPVA 873  
Db 510 TEVLRNLSLARVMDVLYREGDGYTVVGAAGGITSLLEPIA 553

RESULT 13  
S68366  
(+)-delta-cadinene synthase isozyme XC14 - Gossypium arboreum  
C:Species: Gossypium arboreum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: S68366  
R;Chen, X.Y.; Chen, Y.; Heinstein, P.; Davison, V.J. Arch. Biochem. Biophys. 324, 255-266, 1995  
A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a cDNA from Gossypium arboreum  
A:Reference number: S68365; MUID:96132653; PMID:8554317  
A:Accession: S68366  
A:Molecule type: mRNA  
A:Residues: 1-554 <CHE>  
A:Cross-references: EMBL:U23205; NID:g1045313; PIDN:AAA93065.1; PID:g1045314  
A:Experimental source: cultivar Nanking  
C:Superfamily: vetispiradiene synthase 1  
C:Keywords: phytoalexin biosynthesis

Query Match 13.9%; Score 643.5; DB 2; Length 554;  
Best Local Similarity 30.8%; Pred. No. 2.5e-36;

QY	475	FTTNHLQNALAKXNNAFDKWAVKQDLPGEEVEYAIKYPWHRSMPLREARSYTEQFGSNDVWL	534
Db	172	FTSIHLESAAPH-----LKSPLEQVTHALEQCCLHKGVPRVETRFYI-----	213
QY	535	GKTVYKMLYVSNBKYLELAKLDFNMVQALHQKETOHIVSWRRSGF-NDLTFTRQRPVEM	593
Db	214	-SSIIDKEQSKNNVLLRFKALDENLLQMLKQBLAQVSRWKKOLDFTVTLLPYADRVRVEC	272
QY	594	YFSVAVSMPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDCLKLFSEAVRRWDISVLDS	653
Db	273	YFWALGVYFEPQYSQARVMLVKIISMTISIVDDTFDAYGTVKLEAYTDAIQRWIDNEIDR	332
QY	654	VRDNQLKVCFLGLYNTVNGFGKGLKQGRD-VLGYLRKRWEGGLLASYTKEAEWSAAKYV	712
Db	333	LPD-YMKISYKAILDLYKDYKE-LSSAGRSHIVCHAIERMKVEYVRNYNVESTWPIEGYM	390
QY	713	PTFNEYVENAKVSTALATVNLNSIFFTGEI-----LPDYILQOVDLRSKFLHLVSLTGRLI	768
Db	391	PPVSEYLSN-----ALATTTTYYIATTSYLGKMSATEQDFEWLSKNPKILEASVICRVI	445
QY	769	NDTKTYQAEARNRGLBSVQCYMRENPECTEELSHVYGIIDNALKELNWEANPASNA	828
Db	446	DDTATYVEKSRGGIATGIECCMEDYDIST-KEAMAKFQNNMAETAWKDINEGLLRP---T	501
QY	829	PLCVRRLL-LPNTARVWOLFMYR-DGFGISDKEMKOHVSRITLDPVP	872
Db	502	PVSTEFLLTPILNARIVEVTYIHNLDGYTHPEKVLKPHIINLLVDSI	548

RESULT 15

T08174

sesquiterpene cyclase (EC 2.5.1.-) - pepper

C:Species: Capsicum annuum (pepper)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: T08174

Risack, K.; Shin, D.H.; He, S.

Plant Cell Physiol. 39, 899-904, 1998

A:Title: Cloning and bacterial expression of sesquiterpene cyclase, a key bran

A:Reference number: Z16395; MUID:99033462; PMID:9816674

A:Accession: T08174

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-559 <BAC>

A:Cross-references: EMBL:AF061285; NID:g3108342; PIDN:AAC61260.1; PID:g3108343

C:Function:

A:Pathway: the synthesis of phytoalexin capsidiol

A:Note: unduced by UV

C:Superfamily: vetesipiradiene synthase 1

C:Keywords: isoprenoid biosynthesis; transferase

Query Match	13.6%;	Score	633.5;	DB 2;	Length	559;			
Best Local Similarity	30.7%;	Pred. No.	1.2e-35;						
Matches	163;	Conservative	107;	Mismatches	202;	Indels	59;	Gaps	155

QY	355	LPRLWAVDVTVERLIGDRYFKKEIKESLDYVYRWDAERGCVGWARCNPIPDVDDTANGLR	414
Db	69	LSETNLIDVIERLGIAYHFEKEIIDLRIYNENSFEGDVYNE-----DLCTCRLOFR	122
QY	415	ILRLHGVNVSVDVLEFRDEKDFCFAGQTQIGVTDNLNLRCISOVCFPGCKIMEPAKT	478
Db	124	LLRQHGYNISUKIFSKFPLDGNRLKSLASDVGL---LSLYEASHVSRHGEDILEDALA	180
QY	475	FTTNHLQNALAKXNNAFDKWAVKQDLPGEEVEYAIKYPWHRSMPLREARSYTEQFGSNDVWL	534
Db	181	FSTHLESATPH-----LEVPLKEQVRHAEQSLHGKIPRIETIOFI-----	222
QY	535	GKTVYKMLYVSNBKYLELAKLDFNMVQALHQKETOHIVSWRRSGF-NDLTFTRQRPVEM	593
Db	223	-SSVYDQAIKNDVLLRFKALDYNMLQMLKQBLAEAVSRWKKOLDINFTLLPYADRVRVEC	281
QY	594	YFSVAVSMPEPEFAACRIAYAKTSCLAVILDDLYDTHGSLDCLKLFSEAVRRWDISVLDS	653

Db 282 YFWALGVYBPOYSQARVMLVKTIAMISIVDDTYDAYGTVDLAIYTDVIQRWDIKEIDS 341  
 Qy 654 VRDNLKVCF---LGLYNTVNGFGKGLKEQGRD-----VLGYLRKVKWEGLLASYTTKEAW 706  
 Db 342 LPD-YMKISYKALLDLY-----KDYEKEMSRDGRSHVYVYAKERLKELVKSYNIEAKW 393  
 Qy 707 SAAKYVPTFNEYVENAKVIA---LATVVLNSIFFTGELLDPVILQQVDLRS---KFLHL 760  
 Db 394 FIEGHPPASEYLRLNAFVTTYTYLATTSYLGMKYAKE-----QQFEWLSKNPKILEG 446  
 Qy 761 VSLTGLINDTKYQABRNRGELVSSVQCYMRENPECTEBEALSHVYGIIDNALKEINWE 820  
 Db 447 CVTICRVIDDIATVEVEKRGQLSTGIECYMRDYSYST-KEAMAKFQEMGESCKDKINEG 505  
 Qy 821 LANPASNAPLCVRRLLFNTRVMOQLFMY-RDGFGISDKEMKOHVSRTLPD 870  
 Db 506 MLRP-TPIPMBFLSRILNLARLVDVTKHNEGYTHPEKVIKPHIIMVVD 555

Search completed: August 17, 2004, 21:03:06  
 Job time : 23.2551 secs

**THIS PAGE IS BLANK**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:56:14 ; Search time 10.7844 Seconds  
(without alignments)  
4215.079 Million cell updates/sec

Title: US-10-041-007-2

Perfect score: 4645

Sequence: 1 MAGVLFANLPCSLQSPKVP.....GISDKEMKHVSRTLFPDPA 873

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1959	42.2	862	1	TASY_TAXBR Q41594 taxus brevi
2	1945	41.9	862	1	TASY_TAXBA Q83va3 taxus bacca
3	1915	41.2	862	1	TASY_TAXCH Q9tf37 taxus chine
4	1194.5	25.7	637	1	TS03_ABIGR Q22340 abies grand
5	1167.5	25.1	627	1	TS02_ABIGR Q24474 abies grand
6	1151.5	24.8	628	1	TS01_ABIGR Q24475 abies grand
7	647.5	13.9	554	1	DCS1_GOSAR Q39761 gossypium a
8	643.5	13.9	554	1	DCS2_GOSAR Q39760 gossypium a
9	643.5	13.9	554	1	DCS4_GOSAR Q49853 gossypium a
10	636.5	13.7	548	1	SEAS_TOBAC Q40577 nicotiana t
11	631.5	13.6	555	1	DCS3_GOSAR Q43714 gossypium a
12	622.5	13.4	554	1	DCS1_GOSHI P93665 gossypium h
13	594.5	12.8	601	1	CASS_RICCO P59287 ricinus com
14	227.5	4.9	516	1	YAKT_RHISN P55538 rhizobium s
15	214.5	4.6	516	1	Y449_ERAJA Q45221 bradyrhizob
16	132	2.8	2324	1	COAC_CHICK P11029 gallus gall
17	126.5	2.7	908	1	CNA3_MOUSE Q80xk6 mus musculu
18	125.5	2.7	2215	1	SORL_MOUSE Q88307 m sortilin-
19	124	2.7	793	1	SYL_MYCPN P75398 mycoplasma
20	123	2.6	1063	1	DPOM_CLAPU P22373 claviceps p
21	122.5	2.6	738	1	SECG_DROME Q9v8k2 drosophila
22	122.5	2.6	757	1	YVDK_BACSU Q06993 bacillus su
23	121	2.6	722	1	CNA3_HUMAN Q66by7 homo sapien
24	121	2.6	861	1	SYL_BUCBP P59433 buchnera ap
25	120.5	2.6	1037	1	N120_YEAST P57429 saccharomyc
26	118.5	2.6	955	1	SYV_BUCAI P57447 buchnera ap
27	118.5	2.6	1064	1	YF08_METJA Q60307 methanococc
28	117.5	2.5	775	1	REP_BPHPI P51711 bacterioph
29	117	2.5	562	1	SVR_THETN Q8r786 thermosacch
30	117	2.5	1075	1	PS72_SCHPO Q13919 schizosacch
31	117	2.5	1634	1	DPOL_METJA Q58295 methanococc
32	117	2.5	2346	1	COAL_BOVIN Q9ct83 bos taurus
33	117	2.5	2346	1	COAL_SHEEP Q28559 ovis aries

#### ALIGNMENTS

##### RESULT 1

ID	TASY_TAXBR	STANDARD;	PRT;	862 AA.
AC	Q41594; Q94FV8;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DB	Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).			
GN	TDCL.			
OS	Taxus brevifolia (Pacific yew).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.			
OX	NCBI_TaxID=46220;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96199163; PubMed=8621577;			
RA	Wildung M.R., Croteau R.B.;			
RT	"A cDNA clone for taxadiene synthase, the diterpene cyclase that			
RT	catalyzes the committed step of taxol biosynthesis.";			
RT	J. Biol. Chem. 271:9201-9204(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21297238; PubMed=11404343;			
RA	Trapp S.C., Croteau R.B.;			
RT	"Genomic organization of plant terpene synthases and molecular			
RT	evolutionary implications.";			
CC	Genetics 158:811-832(2001).			
CC	-!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid			
CC	intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the			
CC	parent olefin with a taxane skeleton.			
CC	-!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +			
CC	diphosphate.			
CC	-!- PATHWAY: Taxol biosynthesis; first step.			
CC	-!- SIMILARITY: Belongs to the terpene synthase family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; U48796; AAC49310.1; --			
DR	EMBL; AF326519; AAK83566.1; --			
DR	HSSP; Q40577; SEAU.			
DR	InterPro; IPR008930; Terp_cyc toroid.			
DR	InterPro; IPR001906; Terp_synth-like.			
DR	InterPro; IPR005630; Terpene synth C.			
DR	InterPro; IPR008949; Terpenoid_synth.			
DR	Pfam; PF01397; Terpene synth; 1.			
DR	Pfam; PF03936; Terpene synth_C; 1.			
KW	Taxol biosynthesis; Lyase.			
FT	ACT_SITE 758 758			BY SIMILARITY.
FT	ACT_SITE 835 835			BY SIMILARITY.
FT	ACT_SITE 839 839			BY SIMILARITY.

Q19020 caenorhabdi  
Q59037 methanococc  
Q9wxh0 homo sapien  
Q9wvf7 mus musculu  
Q9ajp6 arabidopsis  
P52977 caulobacter  
Q1a29 b replicase  
Q9v439 b replicase  
P38605 arabidopsis  
Q9v6j0 homo sapien  
Q00798 plasmodium  
Q8nf91 homo sapien

34 116.5 2.5 4568 1 DYHC\_CABEL.  
35 114.5 2.5 1169 1 SMC\_METJA  
36 114.5 2.5 6885 1 SNE2\_HUMAN  
37 114 2.5 2283 1 DPOE\_MOUSE  
38 113.5 2.4 440 1 FUL0\_ARATH  
39 113.5 2.4 799 1 LON\_CAUCR  
40 113.5 2.4 7094 1 R1AB\_CVSEN  
41 113.5 2.4 7094 1 R1AB\_CVBLU  
42 113 2.4 759 1 CASI\_ARATH  
43 113 2.4 2220 1 CABI\_HUMAN  
44 112.5 2.4 2869 1 RBPI\_PLAVB  
45 112.5 2.4 8797 1 SNE1\_HUMAN

Thu Aug 26 08:27:52 2004

```

FT CONFLICT 148 148 L -> V (IN REF. 2).
FT CONFLICT 767 767 A -> V (IN REF. 2).
SQ SEQUENCE 862 AA; 9803 MW; 9141B59780CD79A1 CRC64;

Query Match
Best Local Similarity 42.2%; Score 1959; DB 1; Length 862;
Matches 385; Conservative 167; Mismatches 255; Indels 76; Gaps 15;

QY 33 KRSSFGNAQH---C-VRSHLRWNC-----VGIHAS 61
DB 11 KKNALGNKAIHDPTNCRAKSERQMMWVCSRSGRTRVMSRSGGPGVVMSSSTGTSKV 70
QY 62 AAEATRP---DQLPOERFVSRNLADYHPAVWKDDFIDSLTSPNHSATSKSSVDETINKRI 118
DB 71 VSETSTIVDDIP-----RLSANYHGDLMHNVQIOTLETPRESS-----TYQERA 116
QY 119 QTLVKIEIQCFMSGMDGTNPSTAYTAWARIPSI--DGSAPQFPOTLQWILNNQLPDG 176
DB 117 DELVVKIKOMFNALGDGDISPSAYDTAWARVATVSSDGEKPRFPQALNWNQLODQ 176
QY 177 SWGECIFLAYDRVLNTLACLLTLKWNKGDIQVQGVFVRKHEEMKDEADNHRPSGF 236
DB 177 SWGIESHFLCDRLNTNSVIALSVKWTGHSQVQGAFFAENLRLLNEE--DELSDF-F 234
QY 237 EVVFPAMLDRAKSLGLDLPYHLPFISQIHQKQKCL-----QKIPLVNLHNTALLY 289
DB 235 QIIFPALLQKAKALGNLPPYDLPFIKYLSTTREARLTDVSAADNIPANMLN----- 286
QY 290 SLEGLQDVVDWQEIITNLSQSDGSLSSPASTACVFMHTQKRCLEHFLNVLKFGDVVPC 349
DB 287 ALEGLEEVDWKNKIMRFSQSGSLSSPASTACVFMHTQKRCLEHFLNVLKFGDVVPC 346
QY 350 HYPDLDFERLWADTVLGRIGDRYFKKEIKESLDVYRYMDAERGVMGWARCNPIPDVDDT 409
DB 347 MYSIDLLERLSVDNIIEHLGIGRHFKEIKVALDVYRYHM--SERGIGWGRSLVPLDNTT 405
QY 410 AMGLRLILRHGVNVSDDVLENFRDEKDFCFAGQTOIGVTDNLNLYRCSQVCFPEKIM 469
DB 406 ALGLRLRMHGVNVSDDVLENFRDEKDFCFAGQTOIGVTDNLNLYRCSQVCFPEKIM 465
QY 470 EEAKTFTTNHQLAKNNAFDKQAVKQDLPGEVEYAIKYPWHRSPRLEARSYIEQFQS 529
DB 466 DDARKFAEYLRALA-----TKISTNKLFEIYVVEYPMHMSIPRLEARSYIDSDYD 520
QY 530 NDVVLGKTVYKMLYVSNKYLEAKLDFNNVQALHOKETQIHVSWHRESGFNDLTTROR 589
DB 521 NYVWQKTLVMPSLNSKCLAKLDFNTVQSLHQBELKLTTRWKEGMDINFTRRH 580
QY 590 PVMEYFSVAVMPEPEPAACRIAAVAKTSCLAVILDDLDYTHGSDLDLKLFEAVRRWDIS 649
DB 581 VAEYVFSAT--FEPEYSATRIAPTKIGCLOVFDMDADIFATLDLKSFTGKRWDT 638
QY 650 VLDSVRDNLKVCPLGLYNTVNGFGDKLKEQGRDVLGYLRKRWEGLLASYTKEAESAA 709
DB 639 LLHIEPE-CHQTCFKWFKLMEEVNDVWVKVQGRDMLAHIRKPMWELFYCYQERWLEA 697
QY 710 KYVPTFNEYENAKVSTALATVLSNIFFTGELLDPYILQOVDLRSKFLHLVSLTGRILN 769
DB 698 GYIPTFEYLKTYAISVGLSPTCTIQLILLMGBELKDDVKEKVPYPSNMFLVSLSWLTN 757
QY 770 DTKYQARNRGELVSSVQCYMRNBPCTEALSHVYGIIDNALKELMELANPASNAP 829
DB 758 DTKYQAEKARGQOASGACIYKONPGATEEDAIIKICRVVDRAKESPEYFPPSNDIP 817
QY 830 LCVRLLLFNTARVQLFWYRDGFGISDKEMKDHVSRFLFDPV 872
DB 818 MGCKSFIEFLNLCVQIFKFDGVIAGNEEIKYIRKVIYIDPI 860

RESULT 2
TASY TAXBA
ID TASY TAXBA
AC Q93YA3;

```

```
Qy 410 AMGLRIILHGYVSSDVLNFRDEKGFPCFAGQTQIGVTDNLNLRCSQVCPFGKIM 469
Db 406 ALGRLTLRTHGYDSSDVLNFRDENGFRFFSAGQTHVELSSVNLFRASDLAPDEGAM 465
Qy 470 BEAKTFTNHLQNALAKNNAFDKAWKDLPGVEYAIKYPWHRSMRPLEARSYIEOFGS 529
Db 466 DDARFAEPYLRDALA-----TKISTNTKLYKEIYVVEYVWMSIPRLEARSYIDSYDD 520
Qy 530 NDVWLGKTVYKMLVSNKYLELAKLDFNMVQALHQKETOHIVSWWRESGFNDLTFTRQR 589
Db 521 DYVWQRTLYRMPSLNSKCLLAKLDFNIVQSLHQBELKLLTRWKEGSMADINFTRHR 580
Qy 590 PVMEYFSAVSMPEPEPAACRIAYAKTSCLAVIILDDLYDTHGSLDLKLFSEAVRRWDIS 649
Db 581 VAEYFSSAT--FEPEYSATRIATFKIGCLQVLFDDMADIPATIDELKSFTEGVKRWDT 638
Qy 650 VLDSDRNOLKVCFLGLVNTVNGFGKGLKEQGRDVLGRLKRWEGLLASYTKAEWSAA 709
Db 639 LLIHPIE-CMOTCFKAVWFKLMEEVNNDVVKQGRDLAHIRKPELFCVQREWLEA 697
Qy 710 KYVPTFNEYVENAKVSIALATVVLNSIPFTGELLPDYILOQVDRSKFLHLVSLTGRLIN 769
Db 698 GYIPTEBYLKYAISVGLGCTLPILLMGLKVDVVEKHYVPSNMFELVLSWRLTN 757
Qy 770 DTKYQARNRNGELVSSVQCYMRNPECTEALSHVYGIIDNALKEINWELANPASNAP 829
Db 758 DTKYQAEKARGOQASGIACVMKONPGATEBDAIKHICRVVDRLKAESEYFEPKSNIP 817
Qy 830 LCVRRLLFNTARVMQFVYMRDGGISDKEMKHVSRTLPDPV 872
Db 818 MGCKSFIFNLRLCQVIFKFDIGYIANEBIKYIRKYIDPI 860

RESULT 3
TASY TAXCH STANDARD; PRT; 862 AA.
AC Q9T37;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).
GN TDCL.
OS Taxus chinensis (Chinese yew).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.
OX NCBI_TaxID=29808;
RN 1;
RP SEQUENCE FROM N.A.
RC TISSUE=Callus;
RA Wang W., Shi Q., Ouyang T., Zhu P., Cheng K.;
RT "Cloning, expression, and characterization of taxadiene synthase, a
RL diterpene cyclase from Taxus chinensis.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid
CC intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the
CC parent olefin with a taxane skeleton.
CC -!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +
CC diphosphate.
CC -!- PATHWAY: Taxol biosynthesis; first step.
CC -!- SIMILARITY: Belongs to the terpene synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY007207; AAG02257.1; --
CC HSSP; Q40577; SEAU.
DR InterPro; IPR008930; Terp_cyc_toroid.
```

```
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene_synth_C.
DR InterPro; IPR008949; Terpenoid_synth.
DR Pfam; PF01397; Terpene_synth_1.
DR Pfam; PF01936; Terpene_synth_C; 1.
KW Taxol biosynthesis; Lyase.
FT ACT_SITE 758 758 BY SIMILARITY.
FT ACT_SITE 835 835 BY SIMILARITY.
FT ACT_SITE 839 839 BY SIMILARITY.
SQ SEQUENCE 862 AA; 98069 MW; 3A597BAF722BF679 CRC64;

Query Match 41.2%; Score 1915; DB 1; Length 862;
Best Local Similarity 45.0%; Pred. No. 4.9e-123;
Matches 370; Conservative 165; Mismatches 245; Indels 42; Gaps 11;

Qy 60 ASAAETRPDQPOEERFVSRNLADYHPAVWKDQFIDSLTSPNSHATS KSSVDETNIRIQ 119
Db 72 SETSSITVIDIP-----RLSANYHCDLWHNVIOQTLETPPRESS-----TYQERAD 117
Qy 120 TLVKEIQCMPSQMGDETNPDSAYDTAWARIPSI--DGSGAPOFPOTLQWILNNQLPDGS 177
Db 118 ELVVKIKDMFNALGDDGISPSAYDTAWARVATISSDSEKPRFPQALNWNVFNQLDQGS 177
Qy 178 WGECEIFLAVDRVLTACLTLTKWNKGDIOVQKGVFVRKHEEMKDEADNHRPSGE 237
Db 178 WGIESHFSLCDRLNTTNSVIALSVWKTGHSQVEQGTETFAENLRLLNEE-DELSPD-FE 235
Qy 238 VVFPAMLDKASLGLDLPYHLPISTQTHQKQKLL-----QKIPLNVLNHNHQTALLYS 290
Db 236 IIFPALLQKAKALGINLPYDLPIKYISTTREALTDSAAADNIPANMLN-----A 287
Qy 291 LEGLDQVVDVHQTINLQSRDGSFLSSPASTACVFMHTQNKRLHFLNVLKSGDYVPCH 350
Db 288 LEGLEEVMDWKIMRFQSKDGSFLSSPASTACVLMNTGDEKCTFLNLLLVKFGCVPCM 347
Qy 351 YPLDLPERLWAVDVERLGIIDRYFKKEIKESLDYVYVYDAERGVGWACNCPDPDQDTA 410
Db 348 YSIDLLERLSLVNIEHLGIRHFKQEKIKVALDYVYRHW--SERGIGMRGSLVPLDNTTA 406
Qy 411 MGLRIILHGYVSSDVLNFRDEKGFPCFAGQTQIGVTDNLNLRCSQVCPFGKIME 470
Db 407 LGLRLTLRTHGYDSSDVLNFRDENGFRFFSAGQTHVELSSVNLFRASDLAPDEGAMD 466
Qy 471 EAKTFTNHLQNALAKNNAFDKAWKDLPGVEYAIKYPWHRSMRPLEARSYIEQPSGN 530
Db 467 DAKKFAEPYLRDALA-----TKISTNTKLYKEIYVVEYVWMSIPRSEARSYIDSYDD 521
Qy 531 DWMLGKTVYKMLVSNKYLELAKLDFNMVQALHQKETOHIVSWWRESGFNDLTFTRQP 590
Db 522 YVWERKTLVYRMPSLNSKCLLAKLDFNIVQSLHQBELKLLTRWKEGSMADINFTRHRV 581
Qy 591 VMEYFSAVSMPEPEPAACRIAYAKTSCLAVIILDDLYDTHGSLDLKLFSEAVRRWDIS 650
Db 582 AEYVFSAT--FEPEYSATRIATFKIGCLQVLFDDMADIPATIDELKSFTEGVKRWDTSL 639
Qy 651 LDSVRDNLKVCFLGLVNTVNGFGKGLKEQGRDVLGRLKRWEGLLASYTKAEWSAAK 710
Db 640 LHEIPE-CMOTCFKAVWFKLMEEVNNDVVKQGRDLAHIRKPELFCVQREWLDAG 698
Qy 711 YVPTFNEYVENAKVSIALATVVLNSIPFTGELLPDYILOQVDRSKFLHLVSLTGRIND 770
Db 699 YIPTFPEYLKYAISVGLGCTLPILLMGLKVDVVEKHYVPSNMFELVLSWRLTND 758
Qy 771 TKYQARNRNGELVSSVQCYMRNPECTEALSHVYGIIDNALKEINWELANPASNAPL 830
Db 759 TKYQAEKARGOQASGIACVMKONPGATEBDAIKHICRVVDRLKAESEYFEPKSNIPM 818
Qy 831 CVRRLLFNTARVMQFVYMRDGGISDKEMKHVSRTLPDPV 872
Db 819 GKCSFIFNLRLCQVIFKFDIGYIANEBIKYIRKYIDPI 860

RESULT 4
```

QY	531	DVWLGKTVYKLVYVNEKYLEAKLDFNMVQALHOKETQHVSWWRESGNDLFTFRQP	590
Db	298	GY---ESLNEMPYNNMKLLQJLAKLEFNIHSLQRELQSIKRWKESGSQLTFTRRH	354
QY	591	VEYFVSVMFPEFAACRIAYAKTSCLAVIILDDLYDTHGSLDDDLKLFSEAVRWDISV	650
Db	355	VEYVWASCIISMLPKHSFAFMFVKVCHLVVLDIYDTFTGWNELQJLFTDAIKRWLST	414
QY	651	LDSVRDQKLVCFGLGNTVNGFGKGLKEQGDVLCYLRKVMWEGLLASYTKEAWSAK	710
Db	415	TRMLPE-YMGVGYMDLYQCINEMVEEAKTQGRDMLNYIQNAWEALFDTFMQEAWS	473
QY	711	YVPTFNEYVENAKVSIALATVVLNSIFFTGELLPDYILOQVLDLASKFLHLVSLTGLND	770
Db	474	YLPTEFYLKNAKVSXGSIATLQPIILTDVPLPDYILOEIDYPSRFNELASSILRGD	533
QY	771	TKTQAEIRNGELVSSVQCYMRENPCTEERASHVYGIIDNALKELNWLNPASAPL	830
Db	534	TRCYKADRGEEASAIKYNKOHPPGSIIEEDALNHINAMISDAIRELNWELLRPD	593
QY	831	CVRELLFNARVMQLFYMYRDGFGISDKEMKHVSRTLFDPA	873
Db	594	SSKHAEDITRAFHVYKYRDGYTVSNNETKNLVKTVLEPLA	636
RESULT 5			
TS2	ABIGR	STANDARD;	PRT; 627 AA.
ID	TS2	ABIGR	STANDARD;
AC	024474;		
DT	28-FEB-2003	(Rel. 41, Created)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	10-OCT-2003	(Rel. 42, Last annotation update)	
DE	Myrcene synthase, chloroplast precursor (EC 4.2.3.15).		
GN	AG2.2.		
OS	Abies grandis (Grand fir).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.		
ON	NCBI_TaxID=46611;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.		
RX	MEDLINE=97413772; PubMed=9268308;		
RA	Bohlmann J., Steele C.L., Croteau R.;		
RT	"Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,		
RT	characterization, and functional expression of myrcene synthase, (-)-		
RT	(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase."		
RL	J. Biol. Chem. 272:21784-21792(1997).		
CC	-1- FUNCTION: Involved in defensive oleoresin formation in conifers in		
CC	response to insect attack or other injury. Involved in monoterpene		
CC	(C10) olefins biosynthesis.		
CC	-1- CATALYTIC ACTIVITY: Geranyl diphosphate = myrcene + diphosphate.		
CC	-1- COFACTOR: Manganese and potassium.		
CC	-1- PATHWAY: Oleoresinosis.		
CC	-1- SUBCELLULAR LOCATION: Chloroplast.		
CC	-1- INDUCTION: By wounding.		
CC	-1- SIMILARITY: Belongs to the terpene synthase family.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	EMBL; AF006193; AAB70907.1; -		
DR	InterPro; IPR008930; Terp_cyc_toroid.		
DR	InterPro; IPR001906; Terp_synth-like.		
DR	InterPro; IPR005630; Terpene synth C.		
DR	InterPro; IPR008949; Terpenoid synth.		
DR	Pfam; PF01397; Terpene synth; 1.		
DR	Pfam; PF03936; Terpene synth C; 1.		
KW	Lyase; Manganese; Transib peptide; Chloroplast.		
FT	TRANSIT	1	637
FT	CHAIN	?	CHLOROPLAST (POTENTIAL).
FT	ACT_SITE	533	533
FT	ACT_SITE	610	610
FT	ACT_SITE	614	614
FT	ACT_SITE	614	614
SQ	SEQUENCE	637 AA;	73477 MW; 8E80CD9DDE886898 CRC64;
Query Match			
Best Local Similarity 41.9%; Pred. No. 5.5e-74;			
Matches 244; Conservative 117; Mismatches 187; Indels 35; Gaps 9;			
QY	300	WQE--ITNLQS-RDGSFLSSPASTAC-----VFMHTONKRLHFLNVLKFGDYVPCHY	351
Db	80	WEDDFIQSLSPYGGSSYSERAEVTVBEVKEMFNSIPNNREL-----FGS-----	124
QY	352	PLDLFERLWADTVRLGIDRYFKKEIKESLDYVYRYWDAERGVGWARCNPIPDVDTAM	411
Db	125	QNDLLRLWVDSIERLGDHFRFQNEIRVALDYVYVWKEGICGGRDSTFPDLNSTAL	184
QY	412	GLRLRLHGVNSDVLNENFDEKGDFFCFAGQTQIGVDN-INLYRCSQVCFEGKIME	470
Db	185	ALRLRLHGVNSDVLNENFDEKGDFFCFAGQTQIGVDN-INLYRCSQVCFEGKIME	244
QY	471	EAKFTTNHLQNALAKNAFDKWA VKKDLPGVEVYAIKYPWHRSMPLREARSYIEQFGSN	530
Db	245	EAEIFSASYLKKVLQK-----IPVSNLSEIEVLEVGWHTNLPLREARNYIEVBQS	297

```
FT TRANSIT 1 2 CHLOROPLAST (POTENTIAL).
FT CHAIN 2 627 MYRCENE SYNTHASE.
FT ACT SITE 523 523 BY SIMILARITY.
FT ACT SITE 600 600 BY SIMILARITY.
FT ACT SITE 604 604 BY SIMILARITY.
SQ SEQUENCE 627 AA; 72478 MW; 2E0DA492E0C971FD CRC64;

Query Match 25.1%; Score 1167.5; DB 1; Length 627;
Best Local Similarity 43.2%; Pred. No. 3.7e-72;
Matches 224; Conservative 104; Mismatches 173; Indels 17; Gaps 5;

QY 354 DLFLRLWAVTVVERLGIDRYFKKEIKESLDVYRYWDAERGVCWACNPIPDVDDTAMGL 413
DB 122 DLMQRLMIVDSVERLGIAHFKNEITSALDYVRYWE-ENGIGCRDSDIVDNLSTALGP 180
QY 414 RILRLHGVNVSVDVLENFRDEKGFPCFAGTQIGVTDNLNLRYCSQVCFPGKEIMERAK 473
DB 181 RTLRLHGVTVSEVLKAFQDQNGQFVCSFGQTGEIRSVNLRYASLAFPGCKVMEBAE 240
QY 474 TPTTNHLQNALAKNNAFDKWA VKDLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSN-DV 532
DB 241 IFTSTYRLKEALQK-----IPVSALSQBIKFPVMEYGWHTNLPRLLEARNYIDTLEKOTSA 293
QY 533 WLKGVTVYKMLYVSNKYLELAKLDFNMVQALHQKETQIHVSWWRSGFNDLFTFRQDVE 592
DB 294 WLKNK-----AGKLELAKLEFNINFSQKQELQVLLRWKESDLPKLTFAHRHVE 346
QY 593 MYFSVAVSMFPEPAACRIAYAKTSCLAVILDDLYDTHGSLDDLLKLFSEAVRRWDISVLD 652
DB 347 FYTLASCIADPKHSAFLGPAKMHVTLVDDIYDTFTGTTIDELELFTSAIKRWNSSIE 406
QY 653 SVRDNLKVCFLGLYNTVNGFGKGLKEQGRDVLGYLRKWEGLLASYTEAEASAKYV 712
DB 407 HLPB-YMKCVTVMVVFETYNELTREAKTQGRNTLNYRKAWEAYPDVDMEEAKMISGYL 465
QY 713 PTFNEYVNAKVSIALATVNLNIPFTGELLPDYILOQVDLRSKFLHLVSLTGLINDTK 772
DB 466 PMFEYHENGKVVSAAYRVATQPILTNAPLDYILKIGIDPSPNDLASSFLRLRGDTR 525
QY 773 TYQARNRGELVSSVQCYMRNPECTEBEALSHVYGIIDNALKEINWELANPASNAPLCV 832
DB 526 CYKADRDGEEASCISCYMKDNGSTEDALNHINAMVNDIINKELNELLNSNDNIPMLA 585
QY 833 RLLPNTARVQLFMYVDGGGIDSKENKDHVSRTLED 870
DB 586 KGHAFDITRALHLLIYIRDGFSVANKETKCLVMETLLE 623

RESULT 6
TSID1 ABIGR STANDARD; PRT; 628 AA.
AC 024475;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pinene synthase, chloroplast precursor (EC 4.2.3.14) (Beta-geraniolene synthase) ((-)-(1S,5S)-pinene synthase).
DE GN AG3.18.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=97413772; PubMed=9268308;
RA Bohlmann J., Steele C.L., Croteau R.;
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase."
RL J. Biol. Chem. 272:21784-21792(1997).
CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in response to insect attack or other injury. Involved in monoterpene (C10) olefins biosynthesis. A mixture of alpha-and beta-pinene is produced by this enzyme.
CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate = pinene + diphosphate.
CC -!- COFACTOR: Manganese and potassium.
CC -!- PATHWAY: Oleoresinosis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- INDUCTION: By wounding.
CC -!- SIMILARITY: Belongs to the terpene synthase family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; U87909; AAB71085.1; --
DR HSPP; Q40577; SEAT.
DR InterPro; IPR008930; Terp_cyc toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR InterPro; IPR005630; Terpene_synth_C.
DR InterPro; IPR008949; Terpenoid_synth.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF01936; Terpene synth; C; 1.
DR Lyase; Manganese; Transit peptide; Chloroplast.
DR TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
FT CHAIN 1 ? PINENE SYNTHASE.
FT ACT SITE 524 524 BY SIMILARITY.
FT ACT SITE 601 601 BY SIMILARITY.
FT ACT SITE 605 605 BY SIMILARITY.
SQ SEQUENCE 628 AA; 71505 MW; 23DBB788F3C6072C CRC64;

Query Match 24.8%; Score 1151.5; DB 1; Length 628;
Best Local Similarity 44.0%; Pred. No. 4.7e-71;
Matches 230; Conservative 97; Mismatches 179; Indels 17; Gaps 6;

QY 352 PL-DLFLRLWAVTVVERLGIDRYFKKEIKESLDVYRYWDAERGVCWACNPIPDVDDTA 410
DB 119 PLNDLIQRLMIVDSLERLGIHRHFKDIBKSALDYVYSYVW-GENGIGCRGSEVVDLNSA 177
QY 411 MGLRLRLHGVNVSVDVLENFRDEKGFPCFAG-QTQIGVTDNLNLRYCSQVCFPGSKIM 469
DB 178 LGLRLRLHGVNVSVDVLENFRDEKGFPCFAG-QTQIGVTDNLNLRYCSQVCFPGSKIM 237
QY 470 EBAKTFTTNHLQNALAKNNAFDKWA VKDLPGEVEYAIKYPWHRSMRPLEARSYIEQFGS 529
DB 238 DEAEIFSTKYLKEALQK-----IPVSLSRIGDVLVEYGWHTYLPRLLEARNYIQVFGQ 290
QY 530 NDVWLKGVTVYKMLYVSNKYLELAKLDFNMVQALHQKETQIHVSWWRSGFNDLFTFRQR 589
DB 291 D-----TENTKSYKVSQKLELAKLEFNIFQSLOKRELSLVRWVKESGFPMTPCRHR 344
QY 590 PVMEYFSVAVSMFPEPAACRIAYAKTSCLAVILDDLYDTHGSLDDLLKLFSEAVRRWDIS 649
DB 345 HVEYITLASCIAIEPQHSQSGFLGFAKTCHLITVLDNDYDTFGTVDELELFTATWKWDPS 404
QY 650 VLDSVRDNLKVCFLGLYNTVNGFGKGLKEQGRDVLGYLRKWEGLLASYTEAEASAA 709
DB 405 SIDCLPE-YMKGVYIAVYDVTNEMAREABEAQRDTLTAREAEAVIDSYMGEARWIAT 463
QY 710 KYUPTFNEYVNAKVSIALATVNLNIPFTGELLPDYILOQVDLRSKFLHLVSLTGLIN 769
DB 464 GYLPSFDEYENGKVCSCGHRISALQPILTWMDIIPFDHILKEVDPPSKLNDLACAILLRG 523
QY 770 DTKTYQARNRGELVSSVQCYMRNPECTEBEALSHVYGIIDNALKEINWELANPASNAP 829
DB 524 DTKCYKADRGEEASCISCYMKDNGSTEDALNHINAMISDVIGKLNELLKPDINVP 583
QY 830 LCVRRLPNTARVQLFMYVDGGGIDSKENKDHVSRTLED 872
DB 584 ISAKGHAFDITARAFHYKYKRDGYSVANVETKSLVTRTLESV 626
```







Thu Aug 26 08:27:52 2004

us-10-041-007-2.rsp

NCBI\_TaxID=4097;  
 [1] SEQUENCE FROM N.A., AND SEQUENCE OF 56-73.  
 RN STRAIN=ev. NK326;  
 RC MEDLINE=93066390; PubMed=1438319;  
 RX Faccini P.J., Chappell J.;  
 RA "Gene family for an elicitor-induced sesquiterpene cyclase in tobacco."  
 RT Proc. Natl. Acad. Sci. U.S.A. 89:11088-11092(1992).  
 RL [2]  
 RN X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS), AND REVISIONS.  
 RP MEDLINE=97442533; PubMed=9295271;  
 RX Starks C.M., Back K., Chappell J., Noel J.P.;  
 RA "Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase."  
 RT Science 277:1815-1820(1997).  
 RL -!- FUNCTION: Catalyzes the cyclization of trans,trans-farnesyl diphosphate (FPP) to the bicyclic intermediate 5-epi-aristolochene, initial step in the conversion of FPP to the sesquiterpenoid antifungal phytoalexin capsidiol.  
 CC -!- CATALYTIC ACTIVITY: Trans,trans-farnesyl diphosphate = aristolochene + diphosphate.  
 CC -!- COFACTOR: Binds 3 magnesium ions per subunit.  
 CC -!- SUBUNIT: Monomer.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- INDUCTION: By fungal elicitor.  
 CC -!- SIMILARITY: Belongs to the terpene synthase family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L04680; AAA19216.1; --  
 DR PIR; T03714; T03714.  
 DR PDB; SEAS; 15-OCT-97.  
 DR PDB; SEAT; 12-NOV-97.  
 DR PDB; SEAU; 08-APR-98.  
 DR InterPro; IPR008930; Terp\_cyc\_toroid.  
 DR InterPro; IPR001906; Terp\_synth-like.  
 DR InterPro; IPR005630; Terpene\_synth\_C.  
 DR InterPro; IPR008949; Terpenoid\_synth.  
 DR Pfam; PF01397; Terpene\_synth; 1.  
 DR Pfam; PF03936; Terpene\_synth\_C; 1.  
 KW Lyase; Magnesium; 3D-structure.  
 FT ACT\_SITE 273 273 PROTON ACCEPTOR.  
 FT ACT\_SITE 444 444  
 FT ACT\_SITE 520 520  
 FT ACT\_SITE 525 525  
 FT CONFLICT 42 42  
 FT CONFLICT 44 44  
 FT CONFLICT 55 55  
 FT CONFLICT 62 62  
 FT CONFLICT 73 73  
 FT CONFLICT 89 89  
 FT CONFLICT 388 388  
 FT HELIX 26 29  
 FT HELIX 36 57  
 FT TURN 58 58  
 FT TURN 60 61  
 FT TURN 64 76  
 FT TURN 77 78  
 FT HELIX 80 83  
 FT HELIX 84 97  
 FT HELIX 104 116  
 FT TURN 117 118  
 FT STRAND 123 129  
 FT STRAND 130 133  
 FT STRAND 132 133  
 FT STRAND 136 136  
 Y -> YYY (IN REF. 1).  
 K -> Q (IN REF. 1).  
 N -> S (IN REF. 1).  
 M -> R (IN REF. 1).  
 T -> I (IN REF. 1).  
 D -> E (IN REF. 1).  
 T -> M (IN REF. 1).

138	142	HELIX	13.7%;	Score 636.5;	DB 1;	Length 548;
144	154	HELIX	Best Local Similarity	28.9%;	Pred. No. 5.9e-36;	
155	156	TURN	Matches 153;	Conservative 119;	Mismatches 204;	Indels 53; Gaps 15;
159	160	TURN				
162	164	HELIX				
165	166	TURN				
167	178	HELIX				
179	181	HELIX				
184	185	TURN				
186	195	HELIX				
199	200	TURN				
203	213	HELIX				
214	214	TURN				
215	217	HELIX				
219	220	TURN				
223	254	HELIX				
255	255	TURN				
256	259	HELIX				
261	262	TURN				
267	277	HELIX				
281	283	HELIX				
284	306	HELIX				
307	307	TURN				
310	322	HELIX				
323	323	TURN				
326	330	HELIX				
333	354	TURN				
355	357	TURN				
359	361	HELIX				
362	385	HELIX				
386	386	TURN				
391	398	HELIX				
399	399	TURN				
400	402	HELIX				
404	413	HELIX				
414	414	TURN				
416	417	TURN				
420	427	HELIX				
428	428	TURN				
431	454	HELIX				
455	456	TURN				
458	459	TURN				
461	469	HELIX				
470	470	TURN				
473	494	HELIX				
495	495	TURN				
503	505	HELIX				
506	506	TURN				
507	519	HELIX				
520	520	TURN				
535	542	HELIX				
543	543	TURN				
548	548	SEQUENCE	62973 MW;	9FE1C59CF1A68BF1	CRC64;	
Query Match	13.7%;	Score 636.5;	DB 1;	Length 548;		
Best Local Similarity	28.9%;	Pred. No. 5.9e-36;				
Matches 153;	Conservative 119;	Mismatches 204;	Indels 53;	Gaps 15;		
353	LDLFLERLMAVDTVVERLGIDRYFKKEIKESLDYVYRYWDAERGVGWACRNIPDQVDDTAMG	412				
62	MKLADTLNLTDTIERLGISYHFEKEIDDLQIYNQ-----NSNCN---DLCTSLAQ	110				
413	LRIRLHGYNVSSDVLNFEDEKDFCFAGQTQIGVTNLTNLYRCSQVCPPGKIMEEA	472				
111	FRLLRQHGFINISPEIFSKFDENGKFKESLASDVLGL---LNLYESHVRTHADLLEDA	167				
473	KTFTTNHLQNALAKNNAFDKWAVKDLPGGEVVAIKYPMHRSMPRLRSRVIEQFGSNDV	532				
168	LAFSTHLESNAAPH-----LKSPLREQVTHALEQCLHKGVPVETREFI-----	211				
533	WLGKTVYKMLYVSNKEYLELAKLDFNNVQALHOKETQHIYVSWRESGF-NDLTFTFRQPV	591				
212	---SSIYDKESQSKNNVLLRFKLDLDFNLLQMLHKLQAVSRKWKDLDFVTLPLYARDVV	268				

QY 592 EMYFSAVMSPEPFAACRIAYAKTSCLAVIDLDLYTHSLDLDLKFSEAVRRWDISVL 651  
 Db 269 ECFYFALGVYFPEQYSOARVNLVXTISMISIVDDTDPAYGTVELEAYTDAIQRWIDNEI 328  
 QY 652 DSVRDNLKVCFLGLYNTVNGFGDKGLKEQGRD-VLGYLRKRWEGGLASYYTKEAWSAAK 710  
 Db 329 DRLPD-YMKYSYKAILDYKDYKE-ISSAGRSHVCHAIERKKEVVNTVNSTWTFIEG 386  
 QY 711 YVPTFNYENAKVSIALATVVLNSIPTGEL-----LPDYLOQVLDLRSKPLHLVSITGR 766  
 Db 387 YTPVPSEYLSN-----ALATYYLYLATTSYLGKMSATEQDFEWSLKNPKILEASVIICR 441  
 QY 767 LINDTKYQAEERNEGELSSVQCYMRENPECTEERALSHTVYGIIDNALKELNWLNPAS 826  
 Db 442 VIDDTATYEKSRGQATGIECMRDYGI-ST-KEAWKTFQNAETAWKDINEGLLRP-- 498  
 QY 827 NAPLCVRL--LFTARVMOLFMYNR-DGFGISDKEMKDHVSRTLFDPV 872  
 Db 499 -TPVSTFELPILMLARIVEVYIHNLDGTHPEKVLKPHIINLLVDSI 546

RESULT 11  
 DCS3 GOSAR STANDARD; PRT; 555 AA.  
 AC Q43714; 1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE (+)-delta-cadinene synthase isozyme A (EC 4.2.3.13) (D-cadinene synthase).  
 GN CAD1-A.  
 OS Gossypium arboreum (Tree cotton).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Malvales; Malvaceae; Gossypium.  
 OX NCBI\_TaxID=29729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nanking;  
 RX MEDLINE=97060798; PubMed=8904844;  
 RA Chen X.-Y., Wang M., Chen Y., Davison V.J., Heinstein P.;  
 RT "Cloning and heterologous expression of a second (+)-delta-cadinene synthase from Gossypium arboreum.";  
 RL J. Nat. Prod. 59:944-951(1996).  
 CC -1- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl diphosphate (FPP) to (+)-delta cadinene.  
 CC -1- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-delta-cadinene + diphosphate.  
 CC -1- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis; first (committed) step.  
 CC -1- SIMILARITY: Belongs to the terpene synthase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC ENBL; X96429; CAA65289.1; -;  
 CC EMBL; U27535; RAB41259.1; -;  
 CC HSSP; Q40577; SEAT.  
 CC InterPro; IPR008930; Terp\_cyc\_toroid.  
 CC InterPro; IPR001906; Terp\_synth-like.  
 CC InterPro; IPR005630; Terpene\_synth\_C.  
 CC InterPro; IPR008949; Terpenoid\_synth.  
 CC Pfam; PF01397; Terpene\_synth\_1.  
 CC Pfam; PF03936; Terpene\_synth\_C; 1.  
 KW Lyase; Multigene family.  
 FT ACT\_SITE 452 452 BY SIMILARITY.  
 ACT\_SITE 528 528 BY SIMILARITY.

FT ACT\_SITE 532 532 BY SIMILARITY.  
 SQ SEQUENCE 555 AA; 64103 MM; 7060C4F9D99412F9 CRC64;  
 Query Match 13.6%; Score 631.5; DB 1; Length 555;  
 Best Local Similarity 30.3%; Pred. No. 1.3e-35;  
 Matches 159; Conservative 104; Mismatches 217; Indels 45; Gaps 12;  
 QY 354 DLPERLWAVTVVERLIGIDYFKKEIKESLDYVTRYM--DAERGVGWACNPIPDVDDTA 410  
 Db 68 DSNQKLPFDIDAVQRLGVSHFEKEIEDELENIYRDTNNNDAD-----TDLYTTA 116  
 QY 411 MGRILRLHGYNVSSDVLENFRDEKGFCCFACQTIQGVTDNLNLVRCQVCPPEKIME 470  
 Db 117 LRRLREHGFDISCDAFNFKPDAGN---FKASLTSDVQGLLELYEASVNRVHGEDILD 173  
 QY 471 EAKTFTTNHLQNALAKNNAPDKWAVKDLPGFEVEYAIKYPWHRSMRPLEARSVIEQFGSN 530  
 Db 174 EASFTTAQLTALP-----TLHPLSEQVGHALKQSIRRLPRVEARNPI----- 219  
 QY 531 DVLGKTVYKMLYVSNKYLELAKLDFNMVQALHOKETQHIHVSWRRESGF-NDLTFTTRQ 589  
 Db 220 -----SIYQDLSEHNKSLIQFAKIDFNLIQLLRKELSEICRWKDLDFTRKLPFADR 273  
 QY 590 PVMEYFSVAVSMEPEFAACRIAYAKTSCLAVIDLDLYTHSLDLDLKFSEAVRRWDIS 649  
 Db 274 VVEGYFWMGVYFPEQYSLGRKMLTKVIAMASIVDDTDSYATYDELIPYTNAIERWDIK 333  
 QY 650 LIDSVRDNLKVCFLGLYNTVNGFGDKGLKEQGRD-VLGYLRKRWEGGLASYYTKEAWSA 708  
 Db 334 CMNQL-PNYMKISYKALLNVYEM-EQLLANQGRQYRVEYAKKAMTLVQALYLLAKWTH 391  
 QY 709 AKYVPTFNYENAKVSIALATVVLNSIPTGELLPDYILOQVLDLRSKPLHLVSLTGRLI 768  
 Db 392 QNYKPTFEPRDNALPTSGVAMLAITAFVGMGSEVITPETFKAASDPKIIKASTIICRFM 451  
 QY 769 NDTKTYQAEERNEGELSSVQCYMRENPECTEERALSHTVYGIIDNALKELNWLNPAS-N 827  
 Db 452 DDIAEHKFNHRREDDCSAIECYMKQY-GVTAQAYNEFNKHIESSKQVNEEFLLKPTMP 510  
 QY 828 APLCVRLLFNTARVMOLFMYNRDGFISDKEMKDHVSRTLFDPV 872  
 Db 511 TPVLCRL--NLARVMDVLYREGDGYTHVQKAAGGITSLLIDPI 553

RESULT 12  
 DCS1 GOSHI STANDARD; PRT; 554 AA.  
 AC P93625;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE (+)-delta-cadinene synthase (EC 4.2.3.13) (D-cadinene synthase).  
 GN CDN1.  
 OS Gossypium hirsutum (Upland cotton).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Malvales; Malvaceae; Gossypium.  
 OX NCBI\_TaxID=3635;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Westburn M;  
 RA Davis E.M., Chen Y.-S., Essenberg M., Pierce M.L.;  
 RT "cDNA sequence of a (+)-delta-cadinene synthase gene induced in Gossypium hirsutum L. by bacterial infection.";  
 RL (In) Plant Gene Register PGR98-040.  
 RN [2]  
 RP PARTIAL SEQUENCE, AND CHARACTERIZATION.  
 RX MEDLINE=96351891; PubMed=8728715;  
 RA Davis E.M., Tsuji J., Davis G.D., Pierce M.L., Essenberg M.;  
 RT "Purification of (+)-delta-cadinene synthase, a sesquiterpene cyclase from bacteria-inoculated cotton foliar tissue.";  
 RL Phytochemistry 41:1047-1055(1996).  
 CC -1- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl

CC	diphosphate (FPP) to (+)-delta cadinene.
CC	-1- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-
CC	delta-cadinene + diphosphate.
CC	-1- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;
CC	first (committed) step.
CC	-1- INDUCTION: By bacterial infection.
CC	-1- SIMILARITY: Belongs to the terpene synthase family.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ).
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
CC	-----
DR	EMBL; U88318; AAC12784.1; --
DR	HSP; Q40577; SEAU.
DR	InterPro; IPR008930; Terp_cyc_toroid.
DR	InterPro; IPR001906; Terp_synth-like.
DR	InterPro; IPR005630; Terpene synth C.
DR	InterPro; IPR008949; Terpenoid synth.
DR	Pfam; PF01397; Terpene synth; 1.
DR	Pfam; PF03936; Terpene synth_C; 1.
KW	Lyase.
FT	ACT_SITE 451 BY SIMILARITY.
FT	ACT_SITE 527 BY SIMILARITY.
FT	ACT_SITE 531 BY SIMILARITY.
FT	SEQUENCE 554 AA; 64019 MW; 8BCC78AD8CA58816 CRC64;
SEQ	-----
Query Match	13.4%; Score 622.5; DB 1; Length 554;
Best Local Similarity	30.2%; Pred. No. 5.4e-35;
Matches 158; Conservative 104; Mismatches 212; Indels 49; Gaps 12;	
QY	358 RLWAVTVRLGIDRYFKEIKESLDYVY-RYMDAERGVGWGCNRPDPVDDTAMGLRL 416
DB	73 KLAFDSVQGLGVSHTFEIEDELENIYHNNDNEN-----DLYTSLRFL 121
QY	417 RLHGVNVDLENFDEKDFCFAGQTQIGVTDN-----LNLRCQVCFGEKIMEA 472
DB	122 REHGFHSCDVFNKFKDEQGNF-----KSSYTSVDRGLLELYQASLYRVHGEDILDEA 174
QY	473 KTTFTNHLQNALAKNAFDKAVKDLPGVEVAIKYPWHRSPLEARSYIRFQGSNDV 532
DB	175 ISFTSNHLSJAV-----SLDHPLESEVSHALKQSIIRRGIPRVEARHYL----- 218
QY	533 WLKTYVYKMLYVGNKEYLELAKLDFNNVQALHOKETQHVIVSWWRSGF-NDLFTFRQPV 591
DB	219 ----SVYQDIESHNKVLLEPAKIDFNNVQLLHRLKELSEISRWNKDLDFQKLPYARDRVV 274
QY	592 EMTFSVAVSWFEPFPAACRYAYAKTSCLAVILDLDYDTHGSLDLDKLFSEAVRRWDISVL 651
DB	275 EGYFWISGVYFEPQYSLGRKMLTKVIAMASIVDDYDSYATYBELIPYTNAIERWDIKCI 334
QY	652 DSVRDNLKVCFLGLVNTVNGFGKGLKEQGRDV-LGYLRKVGEGLIASLYKEAWSAK 710
DB	335 DELPE-YMKESYKALLDVEEM-EQLVAEHRQYRVEYAKNAMRLAQSILVEARWTLQN 392
QY	711 YVPTFNEYVENAKVSIALATVNLNSIFFTGELIPDYILQVDDLRKSLHLVSLTGLRLND 770
DB	393 YKPSFEFEKANALPTCYGAMLAITSFVGMGDIVTPETFKWAANDPKIIQASTIICRMDD 452
QY	771 TKTYQAEARNRGELVSSVQCVRNPECTEALSHVYGIIDNALKEINWELANPASNAPL 830
DB	453 VTEHKFKHRRHEDDCSAIECYMEY-GVTAQEAQVDFVFNKHVESAWKQVNVQGLFKP-TEMPT 510
QY	831 CVRRLLENTARVMOLFMYRDFGFGISDKMKDHVSRTLFDVPA 873
DB	511 EVLNRSNLARVMDVLYREGDGYTVVGKAAKGITSLIEBIA 553
RESULT 13	
CASS_RICCO	



Search completed: August 17, 2004, 21:00:41  
Job time : 13.7844 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:56:44 ; Search time 46.9123 Seconds  
(without alignments)  
5871.540 Million cell updates/sec

Title: US-10-041-007-2  
Perfect score: 4645  
Sequence: 1 MAGVLPANLPCSLQSPKVP.....GISDKEMKHVSRILFDPVA 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTREMBL 25:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organalle:\*
- 10: sp\_phase:\*
- 11: sp\_plant:\*
- 12: sp\_ricent:\*
- 13: sp\_virus:\*
- 14: sp\_virtebrate:\*
- 15: sp\_unclassified:\*
- 16: sp\_virus:\*
- 17: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4645	100.0	873	10	Q947C4	Q947C4 ginkgo bilo
2	2844	61.2	868	10	Q38710	Q38710 abies grand
3	2822	60.8	853	10	Q94FW1	Q94fw1 abies grand
4	2114.5	45.5	817	10	Q81086	Q81086 abies grand
5	2103.5	45.3	816	10	Q94FW2	Q94fw2 abies grand
6	2085.5	44.9	782	10	Q9SAU6	Q9sau6 abies grand
7	1411	30.4	787	10	Q22667	Q22667 stevia reba
8	1381.5	29.7	800	10	Q9ST35	Q9st35 lycopersico
9	1354.5	29.2	827	10	Q9LRE1	Q9lrl1 scoparia du
10	1352	29.1	829	10	Q9MAX2	Q9max2 croton subl
11	1332	28.7	802	10	Q38802	Q38802 arabidopsis
12	1326.5	28.6	799	10	Q9FXV9	Q9fxv9 lactuca sat
13	1323.5	28.5	823	10	Q9ZTN8	Q9ztn8 cucurbita m
14	1319	28.4	822	10	Q84LM0	Q84lm0 cucurbita m
15	1318	28.4	823	10	Q41771	Q41771 zea mays (m
16	1309.5	28.2	827	10	Q9ZTN7	Q9ztn7 cucurbita m

17	1303	28.1	801	10	O04408	O04408 pisum sativ
18	1257.5	27.1	577	10	O94KA3	O94ka3 picea abies
19	1252	27.0	581	10	O64404	O64404 abies grand
20	1244	26.8	579	10	O94FW3	O94fw3 abies grand
21	1215.5	26.2	630	10	Q9M7D1	Q9m7d1 abies grand
22	1209.5	26.0	634	10	O94KA5	O94ka5 picea abies
23	1203	25.9	637	10	O94FW9	O94fw9 abies grand
24	1193	25.7	637	10	Q9M7C9	Q9m7c9 abies grand
25	1182	25.4	629	10	O94KLA6	O94kla6 pinus taeda
26	1181.5	25.4	633	10	O94KA4	O94ka4 picea abies
27	1169	25.2	628	10	O84KLA3	O84kla3 pinus taeda
28	1151.5	24.8	593	10	O64405	O64405 abies grand
29	1143.5	24.6	785	10	O9FXK5	O9fxk5 cucumis sat
30	1139.5	24.5	623	10	O94FW0	O94fw0 abies grand
31	1139.5	24.5	788	10	O9FXV8	O9fxv8 lactuca sat
32	1135	24.4	685	10	Q7XNG0	Q7xng0 oryza sativ
33	1131	24.3	615	10	O84KLA2	O84kla2 pinus taeda
34	1131	24.3	618	10	O948Z0	O948z0 abies grand
35	1126.5	24.3	627	10	O84SM8	O84sm8 picea abies
36	1116	24.0	630	10	Q9M7D0	Q9m7d0 abies grand
37	1106.5	23.8	627	10	O84KLA4	O84kla4 pinus taeda
38	1095	23.6	784	10	Q9XEH9	Q9xeh9 stevia reba
39	1094.5	23.6	574	10	O84KLA5	O84kla5 pinus taeda
40	1075	23.1	784	10	O9XEI0	O9xei0 stevia reba
41	1072.5	23.1	789	10	Q39548	Q39548 cucurbita m
42	1034	22.3	785	10	O64952	O64952 arabidopsis
43	1033	22.2	785	10	Q9SAK2	Q9sak2 arabidopsis
44	973.5	21.0	830	10	Q852S2	Q852s2 oryza sativ
45	847	18.2	683	10	Q7XLE0	Q7xle0 oryza sativ

ALIGNMENTS

RESULT 1

Q947C4					
ID	Q947C4	PRELIMINARY;	PRT;	873	AA.
AC	Q947C4;				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DS	Levopimaradiene synthase.				
OS	Ginkgo biloba (Ginkgo).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.				
OX	NCBI_TaxID=3311;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21381393; PubMed=11488601;				
RA	Schepmann H.G., Pang J., Matsuda S.P.T.;				
RT	"Cloning and characterization of Ginkgo biloba levopimaradiene				
RT	synthase which catalyzes the first committed step in ginkgolide				
RT	biosynthesis."				
RL	Arch. Biochem. Biophys. 392:263-269 (2001).				
DR	EMBL; AF331704; AL099651; -				
DR	GO; GO:001629; P:lyase activity; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	InterPro; IPR005630; Terpene synth.C.				
DR	InterPro; IPR008949; Terpenoid synth.				
DR	InterPro; IPR008930; Terp_cyc_toroid.				
DR	InterPro; IPR001906; Terp_synth-like.				
DR	Pfam; PF01397; Terpene synth; 1.				
DR	Pfam; PF03936; Terpene synth C; 1.				
SQ	SEQUENCE 873 AA; 100290 MW; P2C9808B19F9931E CRC64;				

Query Match 100.0%; Score 4645; DB 10; Length 873;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MAGVLPANLPCSLQSPKVPFROSTNLIIPFKRSSFGFNAQHCVRSHLRNWCVIHA	60
Db	1	MAGVLPANLPCSLQSPKVPFROSTNLIIPFKRSSFGFNAQHCVRSHLRNWCVIHA	60

RA	Stoffer-Vogel B., Wildung M.R., Vogel G., Croteau R.B.;
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
DR	EMBL; U50768; AAB05407.1; -;
DR	HSSP; Q40577; SEAS.
DR	GO; GO:0016829; P-lyase activity; IEA.
DR	GO; GO:0008522; F-metabolism; IEA.
DR	InterPro; IPR001330; Prenyltrans.
DR	InterPro; IPR005630; Terpene synth C.
DR	InterPro; IPR008949; Terpenoid synth.
DR	InterPro; IPR008930; Terp cyc toroid.
DR	InterPro; IPR001906; Terp synth-like.
DR	Pfam; PF00432; prenyltrans; 1..
DR	Pfam; PF01397; Terpene synth. 1.
DR	Pfam; PF03936; Terpene synth. C; 1.
SQ	SEQUENCE 868 AA; 99536 MW; AD5E79F56B70D25C CRC64;

  

Query Match	61.2%; Score 2844; DB 10; Length 868;
Best Local Similarity	62.3%; Pred.No. 4.1e-200;
Matches 546; Conservative 120; Mismatches 183; Indels 28; Gaps 11;	

  

QY	12	SLQLSPKVP---FRQSTNLIILPHFKKSSGFGNAQHCVSRSHLRLRWN-----CVGIHA 60
Db	5	SSLSQQIPTAAHHLTANAQSIPHFSTTLNAGSSAKRSRLYLWGKSGNKLIACVG--E 62
QY	61	SAATRDPQLPOERFSVR--LNADYHPAVMKDDFDISLTSPNSHATSKSSVDETIKR 117
Db	63	GGATSVPPQSAEKNDLSSTLVKREPPFGWKDDLDLSLT--SHKVAAS--DE---KR 115
QY	118	IQTIVKEITQCWFQSMGDGETNPSTAYDTAWARIPSIDGSGAPQPOTLOLLNNQLPDGS 177
Db	116	LETLISEIKMFRCMGYGETNPSTAYDTAWARIPAVDGSDPNHPETVEVILQNQDKGS 175
QY	178	WGEBCIFLAYDRVLTACLITLKWNKGDIQVKGVEFYRKHMEEMKDEADNRPSGFE 237
Db	176	WGEGFYFLAYDRILLATLCIITLTMTGETQVQKGIEFFRTQAGKMEDEADSHRPSGFE 235
QY	238	VYFPAMLBDAKSIGLIDLPHYLPFTSQIHQRKKLQKIPLNLVHNHOTALLYSLGLEODV 297
Db	236	IYFPAMLKEALIGLDLPYDLPFLLKIIKEKEAKLKRIPTDVLVALPTTLLYSLEGLOEI 295
QY	298	VDWQEITNLQSRDGSFLSPASTACTVFHTONKECLHFLNVLTSKFGDYVPCHVPLDLFE 357
Db	296	VDWQIMKLQSKDGSFLSPASTAAVFNRTGNKKCLDFLNFVLKFGNHVPCHVPLDLFE 355
QY	358	RLWAVDVTVBERLGIDRVYEKKEIKESLDYYRYWDAERGVMGARCNPIPDVDDTAMGLRIIR 417
Db	356	RLWAVDVTVBERLGIDRHFKEEIKEALDYVYSHWD--ERGIGWARENPVPDIDDVTAMGLRIIR 414
QY	418	LHGYNVSSDVLNFRDEKGDFFCFPAGOTQIGVTDNLNLYRCVOVCFCPEKIMEBAKTFTT 477
Db	415	LHGYNVSSDVLKTRFDENGERBEFFCLGTQGRGVTDMLNVCNHSVSFPGETIMEBAKLCTE 474
QY	478	NHLQNALAKNPAFDKWAVKCOLPGEVEVAIKYPWHRSMRPLEARSYIEBFQGSNDVWLCKT 537
Db	475	RYLRLALENDVADFKNWAFKKNIRGEVEYALKYPWKSMPRLEARSYIENYGPDVWLCKT 534
QY	538	VYKMLVYSNEKYLELAKLDFNMVQALHQKETQHIYSWMRESGFNDLTTRQRPVEMYFSV 597
Db	535	VYMPPYISNEKYLELAKLDFNKVQS IHQTELQOLRRWWKSGFTDLNTRFREVTYIYFSP 594
QY	598	AVSMFEPEAPACRIAIAKTSCLAVILDDLTYDHGSLDDLKLFSEAVRWMDISVDSVRDN 657
Db	595	ASPIFEPEFKCREVYTKTSNFTVIILDDLYDAHGS LDDLKLFSTVKRWDLISVDM-PQ 653
QY	658	QLKVCYFGLNTVNGFGKGLKEQGDRDLGLYRKWEGLLASYTKEAWSAAKYVPTTFNE 717
Db	654	QMKITCFGVNYTFNMDIAKEREQGRDVLGIIQNVKVKOLEATYKEAWSAAKYVPSFNE 713
QY	718	YVENAKYSIALATVVVLSNIFFTGELLIPDYILOQVDLRSEFELHVLSITGLINDTKTYQAE 777
Db	714	YIENASVISALGTVVVLSALTFTGEVLTKDEVLSKIDRESFLQMGUIGRUVINDTKTYQAE 773
QY	778	NNRGELVSSVOCYNNRENPECTEBEALSHVYGIIDNALKELNWELANPASNAPLCVRRLLF 837



```

Db 774 RGQGEVASAQCWKDHPKLSBEALQHVSVNENALSELNREFVN--NKIPDIYKRLVP 831
QY 838 NTAHVWQLFYMYRDGFGIS-DKEMKHVSRTLDPVA 873
Db 832 ETARIMQLFYMQGDLTLSDMBEIKHVKNCLEFQVPA 868

RESULT 3
Q94FW1 PRELIMINARY; PRT; 853 AA.
AC Q94FW1;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Abietadiene synthase (Fragment).
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21297238; PubMed=11404343;
RA Trapp S.C., Croteau R.B.;
RT "Genomic organization of plant terpene synthases and molecular
RT evolutionary implications."
RL Genetics 158:811-832(2001).
DR EMBL; AF326516; AAK83563.1; -.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR005630; Terpene synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF00432; Prenyltrans; 1.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
FT NON TER
FT 1
SQ SEQUENCE 853 AA; 98059 MW; 7ECB56123F12B8F1 CRC64;

Query Match 60.8%; Score 2822; DB 10; Length 853;
Best Local Similarity 62.8%; Pred. No. 1.6e-198;
Matches 538; Conservative 121; Mismatches 172; Indels 26; Gaps 11;

QY 29 IPFHKSSGFGNAQHCVSRHLRWN-----CVGIHSAATRPQQLPQBERFVSR- 79
Db 11 IP-HFSTLWAGSSAKRSRLYLWKGSKNIIACVG--EGGATSVPYQSAEKNDLSLS 67
QY 80 --LNADYHPAVWKDDFDLSLTSNSHATSKSVDETINKRIQTLVKEIQCMQSGDGET 137
Db 68 TLVKREFPFGFKDLDLSLTS--SHKVAAS--DE---KRIETLISEIKMFCMGYGET 120
QY 138 NPSAYDTAWARIPSIDGSCAPQPTQLOWLNNQLPDGSGWGECEIFLAYDVLNLIACL 197
Db 121 NPSAYDTAWARIPAVDGSNPHFETVEWILQNLKDGSGWGGFVFLAYDRILATLACI 180
QY 198 LTLKINWKGDIQVKGVEFVRKMEEMKOBADNHRPSGFVFPVPMILDEAKSLGLDLPYH 257
Db 181 IITLWRTGTQVKGIESFRTQAGWDEADSHRPSGFVFPVPMILKEAKILGLDLPYD 240
QY 258 LPFISQIHKQKKLQKIPUNLVNHNQTALLYSLEGLQDVVDWQOETNLSQSDGSLSSP 317
Db 241 LPFLKQIIEKREAKLRIPDVLVYALPTLLYSLEGLQEIWEKIMKLQSDGSLSSP 300
QY 318 ASTACVPMHTQNRCLHFLNVLKSGDYVPCHYPLDLFERLAVDTVERLGDYFKE 377
Db 301 ASTAAVPMRTGNKKCLDLNVLVKFGNHFVCHYPLDLFERLAVDTVERLGDYFKE 360
QY 378 IKESLDVYVRYDAERGQVWARCNPIPDVDDTAMGLRILRLHGVNYSVSSDVLNFRDEKGD 437
Db 361 IKAALDYVYSHWD-ERGIQHARENVPDIDDTAMGLRILRLHGVNYSVSSDVLNFRDENGE 419

```

```

QY 438 FFCFAGOTQIGYTDNLNLYRCSCVCPGGEKIMEEAKFTTNHLONALAKNNAEDKWA VKK 497
Db 420 FFCFLGQTQGVGYDMLNVRNRCVSHVSPGETIMEEAKLTERTYLRNLENADVDFKWA FKK 479
QY 498 DLPGEVEYAIKYVPHRSMRPLEARSYIEQFGSNDVLMGKTYYKMLVYSNEKYLELAKLDF 557
Db 480 NIRGEVEYALKYVPHKSMRPLEARSYIENYGPDDVLMGKTYYVMPYISNEKYLELAKLDF 539
QY 558 NMVQALHQKETOHI VSWWRRESGFNDLTFTRQRPVEMYFVSVAVMFPBPAPACRIAYAKTS 617
Db 540 NKLQSIHQTELQDLRRWKKSSGFTLNFTRRVTETIYFSPASFIFFPEFSKREVTYKTS 599
QY 618 CLAVILDDLADYTHGSLDDLKLPSEAVRWDDISVLDSVRDNLQKVCFLGLNTVNGFGKDG 677
Db 600 NFTVILDDLDYDAHGSLLDKLPFESVWRWDLSDVQDM--PKMKICPVGFTNFNDIAKEG 658
QY 678 LKEQGRDLVGLYLRKVMWEGLLASYTEAEWSAAKYVPTFNEYVENAKVSIATATVNLNISIF 737
Db 659 RERQGRDLVGLYQNVWVKVQLKLEATYKEAEWSEAKYVFSFNEYIENASVSIAGTIVLISAL 718
QY 738 FTGELLDPYILOQVDLRSKPLHLVSLTGRINDTKTYQAEARNRGELVSSVQCYMRENPEC 797
Db 719 FTGEVLTDVLSKIDRESFRLQMLGLTGRLVNDTKTYQAEARGQGEVASAICQCYMKOHPKI 778
QY 798 TEEEAALSHVYGIIDNALKEINWELANPASNAPLCVRELLFNTARVMQLFYMYRDGFGIS- 856
Db 779 SEEAALQHVSVNENALSELNREFVN--NKIPDIYKRLVFETARIMQLFYMQGDLTLSH 836
QY 857 DKEMKHVSRTLDPVA 873
Db 837 DMEIKHVKNCLEFQVPA 853

RESULT 4
ID 081086 PRELIMINARY; PRT; 817 AA.
AC 081086;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE E-alpha-bisabolene synthase.
OS Abies grandis (Grand fir).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.
OX NCBI_TaxID=46611;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98284004; PubMed=9618485;
RA Bohlmann J., Crock J., Jetter R., Croteau R.;
RT "Terpenoid-based defenses in conifers: cDNA cloning, characterization,
RT and functional expression of wound-inducible (E)-alpha-bisabolene
RT synthase from grand fir (Abies grandis).";
RL Proc. Natl. Acad. Sci. U.S.A. 95:6756-6761(1998).
DR EMBL; AF006195; AAC24192.1; -.
DR HSP; Q40577; SEAS.
DR GO; GO:0016829; F:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth; 2.
DR Pfam; PF03936; Terpene synth_C; 1.
SQ SEQUENCE 817 AA; 93749 MW; 95PB06DBC0DE1B4B CRC64;

Query Match 45.5%; Score 2114.5; DB 10; Length 817;
Best Local Similarity 49.5%; Pred. No. 1.6e-146;
Matches 397; Conservative 159; Mismatches 227; Indels 19; Gaps 9;

QY 79 RLNADYHPAVWKDDFDLSLTSNSHATSKSVDETINKRIQTLVKEIQCMFQS--MGDGE 136
Db 25 RRTANPHPNVWGYDLVHSLKSP-----YIDSSYRERAEVLVSEIKMLNPAITGDGE 76

```





252 LDLPVHLPIFQIHOXKOKKLOKIPNLVNLHNTALYSLGSLQVDDVDMQEIITNLQSRDG 311  
 222 IEVPEDETPALKEIYARRDIKTKIPMEVLHVKVPTTLHSLGMPD-LEWEKLLKQCKDG 280  
 312 SFLSSPASTACTVMTQKRCILHFLNLSKFGDYVCHYPLDLPERLWAVDTVERLGID 371  
 281 SFLSPSTAFALMOTKDEKLOYLTNTVTKENGVPVYVVDLFEHIIWVDRQLGLIA 340  
 372 RYFKEIKESLDYVRYWDAERGWCARCNPIPOVDDTAMGLRIURLHGYNVSSVLENF 431  
 341 RYFKEIKDCVEYINKY-TKNGICWARNTHVQDIDDDTAMGPRVLRAHGYDVTDFRQF 399  
 432 RDEKDFCFAGOTQIGVTDNLNLYRCQVCPGKIMEEAKTFTTNLONALAKNAPD 491  
 400 -EKDGKFCVCFAGOSTQAVTGMFNVYRASQMLPFGGERILEDAKKFSYNLYKEKQSTNELLD 458  
 492 KWAVKDLPGVEYEAIKYPMHRSMPRLARSYIEQF-GSNDVWLKTYVYKMLYVSNKYL 550  
 459 KWIIAKDLPGVEGYALDIPWYASLPRLETRYYLEQYGGEDDVMIGTKLYRMGYVSNNTYL 518  
 551 ELAKDLFNNVQALHOKETQHIWVWRESGFNDLTTRQPVEM-YFSVAVSNFEPFAAC 609  
 519 ENAKLDYNNYVAVLOLEWYTIQOYVYDYGIEKFPESDNKSVLYLAASIFEPERSKE 578  
 610 RIAYAKTSCLAVILDDLYD-THGSLDDLKLFSEAVRRWDISVLDSVRDNLKVCFLGLYN 668  
 579 RIWAKTTLIVDKTISFPOSSQSKEDITAFIDKFNKSSKSHSINGEPWHEVWVALCK 638  
 669 -TVNGFGKGLKEQGRDVLGRLKRWEGLLASYTKSAEWSAAKYVPTFNBYENAKVSIAT 728  
 639 TLHGFPALDMTHSDIHPQLHQAEMWMLTKLDQGVDTAELWQVIN-MTAGRWVYKEL 697  
 729 AT-----VNLNIFFTGELLDPYILQOVLDRSKFLHLSLTGLINDT 771  
 698 LTHPOYQLSTVNTSVCHDITLHNPENSTTVDSKVQLVQL---VFSDT 745

RESULT 8  
 ID Q9ST35 PRELIMINARY; PRT; 800 AA.  
 AC Q9ST35;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Copalyl diphosphate synthase.  
 GN CFS.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Solanum.  
 OC NCBI\_TaxID=4081;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Imai R.;  
 RT "Tomato cDNA for copalyl diphosphate synthase."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB015675; BAA84918.1; -  
 DR GO; GO:0016829; P:lyase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR005630; Terpenoid synth. C.  
 DR InterPro; IPR008949; Terpenoid synth.  
 DR InterPro; IPR008930; Terp. cyc. toroid.  
 DR InterPro; IPR001906; Terp. synth-like.  
 DR Pfam; PF01397; Terpene synth.; 1.  
 DR Pfam; PF03936; Terpene synth. C; 1.  
 SQ SEQUENCE 800 AA; 91698 MW; 040E522FE19902F9 CRC64;

Query Match 29.7%; Score 1381.5; DB 10; Length 800;  
 Best Local Similarity 36.6%; Pred. No. 1.2e-92;  
 Matches 312; Conservative 141; Mismatches 275; Indels 125; Gaps 22;  
 80 LNADYHPAVKMDPFIDSLTSPNS-----HATSK----- 107  
 ||| |

12 LTAHQPS-----PSSSPNPQPFKFLKSNREHVFNRILQCHAVSRRRTKDYKEVOSG 64  
 108 -----SSVDETINKRIQTU-----VKBIQCFQSMGMDGETNPAYDTAWVAR 149  
 65 SLPVKWDIDAEVEDETHILEVDPSSNEHDHAIIRSLGSMGDEISVAYDITAWAM 124  
 150 IPSIDSGAPQFPQTLQWILANNQLPDGSGWSECTIFLAYDRVNLTLACLLTKLWNGDIO 209  
 125 VKDVGKTETPQPPSSLEWIANQLADGSGNSIFLVYDRVINTLACIALKSMNLHPDK 184  
 210 VQKGVFVRKMEEMKDEADNHRSGFEVPMMLDEAKSLGLDLPYHLPFISQIHOXK 269  
 185 ILLGMSFRENLSRIGDENAEHMPIGFEVAPSLFEIATKGLDLPYOSVPLQDIYASRQ 244  
 270 KKLQKIPNLVNLHNTALYSLGSLQVDDVDMQEIITNLQSRDGSLSSPASTACVPMHTQN 329  
 245 LKLTRIPKIDIMHVPYTLHSLGMD-LDNQKLLQFQCTDGSFLSPSTAYALMQTD 303  
 330 KRCLEHFLNVLKFGDYVCHYPLDLPERLWAVDTVERLGIDRYFKKEIKESLDYVRYW 389  
 304 HCNLYLKNVAVHKFNGGVENVYVVDLFEHIIWVDRQLGLISRYFELKIKKCIDYFSKYW 363  
 390 DAERGVCWACNPIPDVDDTAMGLRIURLHGYNVSSDVLNENPRDEKGFPCFAGOTQIGV 449  
 364 -TNKGICWARNSPVQDIDDTAMAFRLRLHGVASADVFKEF-ESKGEFFCFVGSQNAV 421  
 450 TDNLNLYRCQVCPGKIMEEAKTFTTNLONALAKNAPDKWAVKDLPGVEYEAIKY 509  
 422 TGMVNLVYASHVWFSGEKILENAKISTSNLYREKRAQNLKDWITKDLPGVEGYALDV 481  
 510 PMHRSMPRLARSYIEQF-GSNDVWLKTYVYKMLYVSNKYLELAKLDENWQALHOKET 568  
 482 PWYASLPRLETRYFLEHYGGEDDVMIGTKLYRMPLVNNLSLYLELAKSYNNCOALHQFEW 541  
 569 QHIVSWRESGFNDLTTRQPVEM-YFSVAVSNFEPFAACRIAYAKTSCLAVILDDLYD 628  
 542 RRIRKWIYECGLREFGLSEKRLVLYYLGSAIFEAQRSTERMAWVKT---AALMDCVRS 598  
 629 THGSLD-DLKLFSEAVRRWDISVLDS--VRDNLKVCFLGLYNTVNGFGKDLKEQGRDV 685  
 599 CFGSPQVSAALFLCFEFAHYSSTALNSRYNTEDRLVGVILG--TLNHLISLALLTHGRDI 655  
 686 LGYLRKRWEGLLASYTK---EAEWSAAKYVPTFNBYENAKVSIATVNLNIFFTGEL 742  
 656 HHYLRHAWENWLLTVGEGEGEGGAELIIRTLN-----LCSVHWISE- 698  
 743 LPDYILQOVLDRSKFLHLSLTGLINDTKTYQ--AERNRGELVSSVQCYTMRENPECTEE 800  
 699 ---EILLS-----HPTYQKLEITNRVSHRLRYKGHSEKQVGMILTFSEE--IEGDMQOLAE 750  
 801 EALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTPARVMQLFY-MYRDFGFGISDKE 859  
 751 LVLSH-----SDASELDANIKDTFTLVAK--SFYYSAYCD-----DRT 786  
 860 MKDHVSRTLPDPV 872  
 787 INFHIAKVLFRV 799

RESULT 9  
 ID Q9LRB1 PRELIMINARY; PRT; 827 AA.  
 AC Q9LRB1;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Copalyl diphosphate.  
 GN CFS.  
 OS Scoparia dulcis.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Plantaginaceae; Gratiolaceae; Scoparia.  
 OC NCBI\_TaxID=107240;

Db 752 VNGGCTWDVGGILSAQIESNQELAKLVPTKSSS---GDLSDTQK----- 794

Qy 831 CVRELLFNTARVMQLFY--MYRDGFGISDKEMKHVSRTLDPV 872

Db 795 -----NFLTARSFYAVY-----CNPGSINFHIAKVLPERV 826

RESULT 10

Q9MAX2

ID Q9MAX2 PRELIMINARY; PRT; 829 AA.

AC Q9MAX2;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Copalyl diphosphate.

GN CPS.

OS Croton sublyratus.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Crotonaceae;

OC Croton.

OX NCBI\_TaxID=107238;

RN [1]

RP SEQUENCE FROM N.A.

RA Sittithaworn W., Sankawa U.;

RT "Copalyl diphosphate synthase from Croton sublyratus.";

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB042424; BAA95612.1; -;

DR GO; GO:0016829; F-lyase activity; IEA.

DR GO; GO:0008152; P-metabolism; IEA.

DR InterPro; IPR005630; Terpene synth C.

DR InterPro; IPR008949; Terpenoid synth.

DR InterPro; IPR008930; Terp\_cyc\_toroid.

DR InterPro; IPR001906; Terp\_synth-like.

DR Pfam; PF01397; Terpene synth; 1.

DR Pfam; PF03936; Terpene synth C; 1.

SQ SEQUENCE 829 AA; 96226 MW; 3C3D00C6A848EFOA CRC64;

Query Match 29.1%; Score 1352; DB 10; Length 829;

Best Local Similarity 37.2%; Pred. No. 1.8e-90;

Matches 293; Conservative 159; Mismatches 250; Indels 86; Gaps 16

Qy 107 KSSVDEITKRIQTLVKEIQCFMSQMGDETNPISAYDTAWARIPSIDGSGAPQPQTILQ 166

Db 105 KLLIKVAISKEMESVERIKQKLSMEDGEINISAYDTAWRLVKDVGSGEPQPSALQ 164

Qy 167 WILNNQLPDGSGWSECIFLAYDVRVLNTLACLTLTKINWKGDIQVQGVFVRKHMEEKD 226

Db 165 WISNNQLPDGSGWDALVFNAHDRILSTLACVVTLKSNWNLPHKVAKGLKFPKENLSKLE 224

Qy 227 EADNHRSGSEVVPAMLDEAKSLGLDPYHLPTFSIQIHQKQKLOKILPLNVLNHQTA 286

Db 225 ENEEHSIGFEVFPSPSLELARKLDIDVPDSDHVLQIQCECRNTKIKKIPKIDIMHKMPTS 284

Qy 287 LLVSLEGLQDVVDQBEITNLSQSDGSLSSPASTACVPMHTQNKRLHFLNVLSKFGDY 346

Db 285 LLHSLEGMFD-LEWEKLLKLFQDNGSFLFSPSSATAYAFMQTDDKCLLYLNLRIVORFNGG 343

Qy 347 VPCHYPDLFLERLWADVTVBERLIGIDRYFKKEIKESLSDYVYRYDWAERGVGMARCNPIPDV 406

Db 344 VPVYVPVDLFEHTWAADRLQRLIGISRLFQEQLKECMNTVAYW-REDGCMCWARNSNVRDV 402

Qy 407 DDTAMGLRIILHGYNYSSVDLENFRDEKGFDFCFAGQTQIGVTNTNLNYRCSQVCFPGE 466

Db 403 DDSSMGFRMLRLYGHNSADYVFKKFKKD-DTWVCMFGQSTQAITGMYNLFRASQVLPFGE 461

Qy 467 KIMEEAKTFTTNHGLNALAKNNAFDKWAVKDLPGEVEYATKYPMHRSMPLREARSYIEQ 526

Db 462 MILEAREFCVDFLREQSANAVVDKMLISKDLPGEVAYALDVPWFASLPTVEARFYIEQ 521

Qy 527 F-GSNDVYMLGKTVYKMLYVSNKYLELAKLDLNNVQALHKQETQHIVSWWRSGFNDLTF 585



DE Copalyl diphosphate synthase No1.  
GN LSCP81.  
OS Lactuca sativa (Garden lettuce).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;  
OC Lactuca.  
OX NCBI\_TaxID=4236;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAINS=Grand Rapids;  
RA Toyomatsu T., Mitsuhashi W., Kamiya Y.;  
RT "Gibberellin biosynthetic enzyme";  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB031204; BAB12440.1; -;  
DR GO; GO:0016829; P:lyase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR005630; Terpene synth C.  
DR InterPro; IPR008949; Terpenoid synth.  
DR InterPro; IPR008930; Terp\_cyc\_toroid.  
DR Pfam; PF01397; Terpene synth; 1.  
DR Pfam; PF03936; Terpene synth C; 1.  
SQ SEQUENCE 799 AA; 91865 MW; 51940A6553B40BA5 CRC64;  
Query Match 28.6%; Score 1326.5; DB 10; Length 799;  
Best Local Similarity 40.8%; Pred. No. 1.3e-88;  
Matches 275; Conservative 126; Mismatches 224; Indels 49; Gaps 14;  
Qy 89 WKDFEID-----SLTSPNSHATSKSVDEITNKRIQTLKEIQCMQSGDGTNPSA 141  
Db 69 WQNDVDELDKCKILYPN-----DE-----IRGFVERIKVLGSMDEGITVSA 114  
Qy 142 YDTAWARIPSIDGSGAPQFPQTOLWLNOLPDGSGEBCIFLAYDRVLNTLACLLTK 201  
Db 115 YDTAWALVQIDNGRFEFPSSLEWIVKNSQSDSGDHLIFSADRIINTLACVIALT 174  
Qy 202 IWNKGDIQVQGVFVRKHMEKDEADNHRPSGFVFPAPMLDEAKSLGLDPLPYHLPFI 261  
Db 175 SWNVHPGCKGLKFLNDNISKLBEENPEHPGFEVAPPSLIDTARKLIDQVPDSPAL 234  
Qy 262 SQIHQKQKLOKIPLVNLHNTALLYSLLEGQVVDVQWQEIITNLSQSDGSLSPASTA 321  
Db 235 KEIYARRNLKTIKPKSLMHKVPPTLLHSLLEGMPD-LEWEKLLKLQKDGSLFSPSTA 293  
Qy 322 CVFHTQNKRCILHFLNVLKSGDVVPCYPLDLFLRLWADTVLRLGIDRYFKKEIKES 381  
Db 294 PALMQTKQKCLQYLTDVATKFNPGVGNVYPVDLFEHIVVDRLQRLGISRYFDSIKDC 353  
Qy 382 LDYVRYVWDAERGVGWACNPIDVDVTAMGLRIILRHGYNVSSDVLNFRDEKGFPCF 441  
Db 354 VDIYRYWTKD-GICWAKNSNVQIDDDTAMGFVLRMHGKYKTTDVPFQF-EKDGKVFVF 411  
Qy 442 AGQTQIGVTDNLNLYRCQVCPGCKIMEAKFTTNHONALAKNAPDKWAVKOLPG 501  
Db 412 PGQTTQAVTGMFNLFRAQSQVLPFDEKILEDAKFSYNYLKEKQSTNELLDKWI 471  
Qy 502 EVEYAIKYPWHRMPLRARSYIEOF-GSNDVWLKTYVKMLYVSNKEYLEAKLDFNVV 560  
Db 472 EVEYALDVPWYASLPRLTRFYLEQGGEDDVWIGKTLRYMGVNSNNTYLEMAKLDYNNC 531  
Qy 561 QALHOKETQHLVSWWRESGF-----NDLTFTQRVPMVYFVAVSMFEPFAACRIAYAK 615  
Db 532 LAIHLEWNTWQOQVDFGMRFGTSDTSL-----LVSYLAAASVFEPEPSKERIAWAK 587  
Qy 616 TSCLAVILDDLYDTHG---SLDDLKLFSEAVRRWDDISVLDSVRDNLQKVCFLGLNTYNG 672  
Db 588 TTTLVDTISSFF--HSLKISNEHREFFVEEFENISIHAKYKPKWGLMVALAGTLHE 645  
Qy 673 FGKGLKQGGQRLVGLYLRKWEGLLASYTK---EAEWSAAKYVPTFN---EYVEN---A 722  
Db 646 IALDVLMTHRDIIHPQLHAWEMLMRWQGVDALEGQAELIVQTNMTAGRWVSNELLA 705

Qy 723 KVSIALATVVLNSI 736  
Db 706 HPQYRLLSVINNI 719  
RESULT 13  
Q9ZTN8  
ID Q9ZTN8 PRELIMINARY; PRT; 823 AA.  
AC Q9ZTN8;  
DT 01-MAY-1999 (TREMELrel. 10, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Copalyl diphosphate synthase 1.  
OS Cucurbita maxima (Pumpkin) (Winter squash).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosidia I; Cucurbitales; Cucurbitaceae; Cucurbita.  
OX NCBI\_TaxID=3661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99063778; PubMed=9847116;  
RA Smith M.W., Yamaguchi S., Ait-Ali T., Kamiya Y.;  
RT "The first step of gibberellin biosynthesis in pumpkin is catalyzed by  
RT at least two copalyl diphosphate synthases encoded by differentially  
RT regulated genes";  
RL Plant Physiol. 118:1411-1419(1998).  
DR EMBL; AF049905; AAD04292.1; -;  
DR GO; GO:0016829; P:lyase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR005630; Terpene synth C.  
DR InterPro; IPR008949; Terpenoid synth.  
DR InterPro; IPR008930; Terp\_cyc\_toroid.  
DR InterPro; IPR001906; Terp\_synth-like.  
DR Pfam; PF01397; Terpene synth; 1.  
DR Pfam; PF03936; Terpene synth C; 1.  
SQ SEQUENCE 823 AA; 93946 MW; 3F367AAP474721F7 CRC64;  
Query Match 28.5%; Score 1323.5; DB 10; Length 823;  
Best Local Similarity 36.8%; Pred. No. 2.2e-88;  
Matches 304; Conservative 148; Mismatches 294; Indels 80; Gaps 22;  
Qy 71 PQERFVRLNADYHPAVWK---DDFIDSLTSPNSHATSKSVDEITNKRIQTLKEIQ 127  
Db 53 PPTQSYSDVLOSQGVPIKQQQFVEDDISETT--AHVLISKEIEERVNR-----IKS 102  
Qy 128 MFQSGDGETNPSAYDTAWARIPSI-DGSGAPQFPQTOLWLNOLPDGSGEBCIFLA 186  
Db 103 MLSSMDDGDISISAYDTAWALIPVLDGVKTPLPFSSLEWIAQNLQPDGSGDGI 162  
Qy 187 YDRVLNTLACLLTLKINWKGDIQVQGVFVRKHMEKDEADNHRPSGFVFPAPMLDE 246  
Db 163 HDRILSTLACVLALNSWKLHPDKSEKGMVFLNKNISKLEDENAEMLLIGFEVAFPSLMF 222  
Qy 247 AKSLGLDPLPYHLPFTISQIHQKQKLOKIPLVNLHNTALLYSLLEGQDV--VDWQEI 304  
Db 223 AKRLNLQVPTDSPVLQEIHNHRSIKLTRIPKEIMHKVPTLLHSLLEGMEGLDWMLL 282  
Qy 305 NLSQSDGSLSPASTACVFMHTQNKRCILHFLNVLKSGDVVPCYPLDLFLRLWADTV 364  
Db 283 KLQAPDGSFLKSPASTAFAMKTNNSCNCFKYLESVVSRENGGVPNVYPVDLFEHIVADR 342  
Qy 365 VERLIGDIYFKKEIKESLDYVRYVWDAERGVGWACNPIDVDVTAMGLRIILRHGYNVS 424  
Db 343 LQRLGSRFFHPHEIVEVDYLRHHW-TDKGICWARDVEFYDDTAMGFKLRLFGHEVS 401  
Qy 425 SDVLENFRDEKGFPCFAGQTQIGVTDNLNLYRCS-QVCFPGCKIMEAKFTTNHONAL 483  
Db 402 AEVFKNF-EKGEFVCYIAGQSTQAVTGMFNLYRASQVWMPFGPKILEDAKQFSYKFLREK 460  
Qy 484 LANNAPFKWAVKQDLPGEVEYAIKYPWHRMPLRARSYIEOF-GSNDVWLKTYVKML 542  
Db 461 QAADELKDKWITTKDLPGEVGVYALDVPWFPASLPVETRYFIQYGGENDIWIQKTLRYMF 520







DR Pfam; PF01397; Terpene synth; 1.  
DR Pfam; PF03936; Terpene synth C; 1.  
SQ SEQUENCE 823 AA; 95123 MW; E4D9C97CDB7A0B34 CRC64;

Query Match 28.4%; Score 1318; DB 10; Length 823;  
Best Local Similarity 35.0%; Pred. No. 5.7e-88;  
Matches 302; Conservative 157; Mismatches 275; Indels 130; Gaps 26;

QY 71 PQERFVSRNLADYHPAVWKDDFDLSLTPNSHATSKSSVDETINK----- 116  
DB 28 PQARRVVERAAAGFGHATTQQ-PDNVSSAKVQTSERVETESKLNRGKPKQDLEDEHQAE 86  
QY 117 --RIQTLVKEIQCMFQSMGDETNPSAYDTAWARIPSI--DGSAPQPPQTLOWILNNQ 172  
DB 87 EABLQPLIDQVRAMLRSMNDGDTASAYDTAWAVPKVGGDGAQPPQFPATVRWIVDHQ 146  
QY 173 LPDGSWGEBCIPLAYRVLTACLTLTKLNKGDIOVQGVFVRKHEEMKDEADNR 232  
DB 147 LPDGSWGDLSALFSDYDRMINTLACVVALTKWLEPARCEAGLGFLENMWRLLAESEASM 206  
QY 233 PSGFEVVPFAMLDKAKSLG-LDLPHYLPFTISOHKROKKLOKIPLNVLNHNHOTALLYSL 291  
DB 207 PIGFEIAPPSLIQTARDLGVDVFPYGHAPALQSIYANREVKKRIPRDMHRVPTSLHSL 266  
QY 292 EGLQDVVDWQIEITNLQSRDGSFLSSPASTACVPMHTQMKRCLHFLNVLKFGDYVPCHY 351  
DB 267 EGMFD-LDWPRLLNLQSCDGSFLFSPSATAYALWQTGDKCFEYIDRIVKKFNGGVENVY 325  
QY 352 PLDLFERLWAVTVRIGIDRYFKKEIKESLDYVYRWDAERGUGVWARCNPIPDVDDTAM 411  
DB 326 PVDLFEHIWVDRLEGLISRYFQRETEQCWDYVNRHW-TEDGICWARKSNVKDVEDDTAM 384  
QY 412 GLRILRLHGVNSVDLENFRDEKDFCFAGOTQIGVTDNLNLYRCOVCFPCKEKIMEE 471  
DB 385 AFRLLRLHGVNSVSPVKNF-EKDGEFFCFVGQQTQAVTGMYNLNRASQISFGEDVLHR 443  
QY 472 AKTFTTNHLQNALAKNNAFDKAVKDLPGEEVYAIKYPWHRSMPLREARSYIEQF-GSN 530  
DB 444 ARVFSYEFLLQREOQGMIRKWIIVAKDLPGEVQYTLDFPFWASLPRVEARTYLDQYGGKD 503  
QY 531 DVMLGKTVYKMLVYSNEKYLELAKLDFNMVQALHOKETQHIVSWRESGFDNLTFTTRQP 590  
DB 504 DVWIGKTLYRMPLVNNDTYLELAIRDFNHCQALHOLECNGLQWYKNCCLDAFGVEPQDV 563  
QY 591 VEMVFSVAVSMFEPFPAACRIAYAKTSLAVILDDLYDTHGSLDDKLKLFSEAVRWDISV 650  
DB 564 LRSYFLAAACIFEPFSAARLAWARTSMIA-----NAISTH-----LRDISDKKRLBCFV 614  
QY 651 LDSVRDNOLK-----VCFLGLYNTVNGFGKGL--KEQGRDVLGLYLRKWEGLLAS 699  
DB 615 HCLYEENDVSWLKNPNPDVILERALRLINLLAQEALPIHEGQRFIHSLLSLAW----- 668  
QY 700 YTKAEAWSAKYVPTFNEYVENAKVSIATATVVLNSIPFTGELLPDYILOQVDLRSKPLH 759  
DB 669 -----TEWMLQKANKENYHKCSGIE-----PQYVWHD---RQTYLL 703  
QY 760 LVSL-----TGRLINDTKTYQAEERNRGELVSVQCYMREN-----ECTEEALSHVYGI 810  
DB 704 LVQVIEICAGRI-----GEAVSMIN--NKDNWFQIQLTCATCDLSNHRMLLS 748  
QY 811 DNALK-----ELNW-----NPASNAPLCVRRLLFNTARVMQLFYMY 849  
DB 749 QDTMKNFARINWIEKEIELNNQELAQSLRLCDEKTSNKK--TKRTLMDVLR--SLYAT 804  
QY 850 RDGFGISDKEMKD-HVSRITLFDV 872  
DB 805 H-----SPQHMIDRHVSRIPEV 823

Search completed: August 17, 2004, 21:02:21  
Job time : 51.9123 secs

THIS PAGE IS BLANK

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: August 24, 2004, 08:52:35 ; Search time 8698.94 Seconds  
(without alignments)  
11165.928 Million cell updates/sec

Title: US-10-041-007-38  
Perfect score: 2241  
Sequence: 1 acgtttcagtcacatgggga.....ttttgatcctgtgcgtag 2241

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : GenEmbl.\*
- 1: gb\_ba.\*
  - 2: gb\_hg.\*
  - 3: gb\_in.\*
  - 4: gb\_om.\*
  - 5: gb\_ov.\*
  - 6: gb\_pat.\*
  - 7: gb\_ph.\*
  - 8: gb\_pl.\*
  - 9: gb\_pr.\*
  - 10: gb\_ro.\*
  - 11: gb\_ste.\*
  - 12: gb\_sy.\*
  - 13: gb\_un.\*
  - 14: gb\_vi.\*
  - 15: em\_ba.\*
  - 16: em\_fun.\*
  - 17: em\_hum.\*
  - 18: em\_in.\*
  - 19: em\_mu.\*
  - 20: em\_om.\*
  - 21: em\_or.\*
  - 22: em\_ov.\*
  - 23: em\_pat.\*
  - 24: em\_ph.\*
  - 25: em\_pl.\*
  - 26: em\_ro.\*
  - 27: em\_ste.\*
  - 28: em\_un.\*
  - 29: em\_vi.\*
  - 30: em\_htg\_hum.\*
  - 31: em\_htg\_inv.\*
  - 32: em\_htg\_other.\*
  - 33: em\_htg\_mus.\*
  - 34: em\_htg\_pln.\*
  - 35: em\_htg\_rod.\*
  - 36: em\_htg\_mam.\*
  - 37: em\_htg\_vrt.\*
  - 38: em\_sy.\*
  - 39: em\_htgo\_hum.\*
  - 40: em\_htgo\_mus.\*
  - 41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2241	100.0	2705	8	AF331704	AF331704 Ginkgo bi
2	1078	48.1	2861	6	BD227676	BD227676 Synthesas
3	1078	48.1	2861	6	AR240711	AR240711 Sequence
4	1078	48.1	2861	6	AR266986	AR266986 Sequence
5	1078	48.1	2861	6	AR316335	AR316335 Sequence
6	1078	48.1	2861	6	AR338478	AR338478 Sequence
7	1078	48.1	2861	6	AR429884	AR429884 Sequence
8	1078	48.1	2861	6	AGU50768	U50768 Abies grand
9	723.4	32.3	2424	6	BD227671	BD227671 Synthesas
10	723.4	32.3	2424	6	AR222107	AR222107 Sequence
11	723.4	32.3	2424	6	AR230276	AR230276 Sequence
12	723.4	32.3	2424	6	AR240706	AR240706 Sequence
13	723.4	32.3	2424	6	AR266981	AR266981 Sequence
14	723.4	32.3	2424	6	AR316330	AR316330 Sequence
15	723.4	32.3	2424	6	AR338473	AR338473 Sequence
16	723.4	32.3	2424	6	AR429879	AR429879 Sequence
17	723.4	32.3	2424	8	AF006194	AF006194 Abies gra
18	723.4	32.3	2525	6	AR230308	AR230308 Sequence
19	723.4	32.3	2528	6	AR230307	AR230307 Sequence
20	723.4	32.3	2528	6	AR230309	AR230309 Sequence
21	723.4	32.3	2528	8	AF006195	AF006195 Abies gra
22	723.4	32.3	2571	6	AR230287	AR230287 Sequence
23	638.2	28.5	2700	6	AR110171	AR110171 Sequence
24	638.2	28.5	2700	6	BD227670	BD227670 Synthesas
25	638.2	28.5	2700	6	AR240705	AR240705 Sequence
26	638.2	28.5	2700	6	AR266980	AR266980 Sequence
27	638.2	28.5	2700	6	AR316329	AR316329 Sequence
28	638.2	28.5	2700	6	AR338472	AR338472 Sequence
29	638.2	28.5	2700	6	AR382645	AR382645 Sequence
30	638.2	28.5	2700	6	AR429878	AR429878 Sequence
31	638.2	28.5	2700	8	TBU48796	U48796 Taxus brevi
32	622.2	27.8	2706	8	AY007207	AY007207 Taxus chi
33	619	27.6	2643	8	AY424738	AY424738 Taxus bac
34	473.2	21.1	2198	8	AF369918	AF369918 Picea abi
35	469	20.9	2150	8	AF369919	AF369919 Picea abi
36	450.4	20.1	2186	8	AR222137	AR222137 Sequence
37	450.4	20.1	2186	8	AF139205	AF139205 Abies gra
38	429.6	19.2	2100	8	AF543530	AF543530 Pinus tae
39	425	19.0	2012	8	AF369920	AF369920 Picea abi
40	417.2	18.6	2186	8	AF461460	AF461460 Picea abi
41	416	18.6	1890	6	AR222146	AR222146 Sequence
42	416	18.6	1893	8	AF139206	AF139206 Abies gra
43	414.2	18.5	2490	8	AB042424	AB042424 Croton su
44	412.2	18.4	2205	6	AR222116	AR222116 Sequence
45	411.2	18.3	1865	6	BD227672	BD227672 Synthesas

ALIGNMENTS

RESULT 1  
AF331704  
LOCUS AF331704  
DEFINITION Ginkgo biloba levopimaradiene synthase mRNA, complete cds.  
ACCESSION AF331704  
VERSION AF331704.1 GI:15865604  
KEYWORDS Ginkgo biloba (maidenhair tree)  
SOURCE Ginkgo biloba  
ORGANISM Ginkgo biloba  
REFERENCE 1 (bases 1 to 2705)  
AUTHORS Schepmann,H.G., Pang,J. and Matsuda,S.P.  
TITLE Cloning and characterization of Ginkgo biloba levopimaradiene synthase which catalyzes the first committed step in ginkgolide



Db 1964 ATCTGTGTGATAGCGTTGGGATTAACAGTTGAAAGTTTGCTTCTAGGGCTGTAC 2023  
 QY 1621 AACACAGTGAATGGATTGGAAAGATGGAAGTCAAGGAAACAGGCCGTGATGCTGGGC 1680  
 Db 2024 AACACAGTGAATGGATTGGAAAGATGGAAGTCAAGGAAACAGGCCGTGATGCTGGGC 2083  
 QY 1681 TATCTTCGAAAGATGGAAGGCTGCTCGCATCGTATACCAAGAGCCGAATGGTCG 1740  
 Db 2084 TATCTTCGAAAGATGGAAGGCTGCTCGCATCGTATACCAAGAGCCGAATGGTCG 2143  
 QY 1741 GCAGCAAGATGTCGCCGACATCAACGAATATGTGAAATGTCGAAGTGCATAGCA 1800  
 Db 2144 GCAGCAAGATGTCGCCGACATCAACGAATATGTGAAATGTCGAAGTGCATAGCA 2203  
 QY 1801 CTTGGCAGTGTGATTAACCTCAATCTTTTCACTGGAGAATTAATCTCTGATTAAT 1860  
 Db 2204 CTTGGCAGTGTGATTAACCTCAATCTTTTCACTGGAGAATTAATCTCTGATTAAT 2263  
 QY 1861 TTACAGCAAGTACGCTTCCGTCGAATTTCTGCATCTTGTCTTTGACTGGACGACTA 1920  
 Db 2264 TTACAGCAAGTACGCTTCCGTCGAATTTCTGCATCTTGTCTTTGACTGGACGACTA 2323  
 QY 1921 ATCAATGACACCAAGACTTACAGGCGGAGAGAAACCGTGTGAAATTTGGTTTCCAGCGTA 1980  
 Db 2324 ATCAATGACACCAAGACTTACAGGCGGAGAGAAACCGTGTGAAATTTGGTTTCCAGCGTA 2383  
 QY 1981 CAGTGTCTACATGAGGGGAAATCCGGAGTGACAGAGGAAGCTCTTAAGTCAATGTTAT 2040  
 Db 2384 CAGTGTCTACATGAGGGGAAATCCGGAGTGACAGAGGAAGCTCTTAAGTCAATGTTAT 2443  
 QY 2041 GGTATCATGACCAACGCACTGAAGGAATGAATTTGGAGTTGGCCAAACCCAGGAGCAAT 2100  
 Db 2444 GGTATCATGACCAACGCACTGAAGGAATGAATTTGGAGTTGGCCAAACCCAGGAGCAAT 2503  
 QY 2101 GCCCCATTGTGTGAGAGACTGCTGTCAACACTGCAAGAGTGAATGCAAGCTGTTTAT 2160  
 Db 2504 GCCCCATTGTGTGAGAGACTGCTGTCAACACTGCAAGAGTGAATGCAAGCTGTTTAT 2563  
 QY 2161 ATGTACAGAGATGGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTGACGCCAACT 2220  
 Db 2564 ATGTACAGAGATGGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTGACGCCAACT 2623  
 QY 2221 CTTTTCGATCTGTGGCGTAG 2241  
 Db 2624 CTTTTCGATCTGTGGCGTAG 2644

RESULT 2  
 LOCUS BD227676 2861 bp DNA linear PAT 17-JUL-2003  
 DEFINITION Syntheses.  
 ACCESSION BD227676  
 VERSION BD227676.1 GI:33037446  
 KEYWORDS JP 2002526066-A/28.  
 SOURCE Abies grandis  
 ORGANISM Abies grandis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.  
 Syntheses  
 Patent: JP 2002526066-A 28 20-AUG-2002;  
 UNIVERSITY OF KENTUCKY RESEARCH DEPARTMENT, THE SALK INSTITUTE FOR  
 BIOLOGICAL STUDIES SECRETARY DEPARTMENT OF HEALTH AND HUMAN  
 SERVICES  
 OS Abies grandis (giant fir)  
 PN JP 2002526066-A/28  
 PD 20-AUG-2002  
 PF 17-SEP-1999 JP 2000574228  
 PR 18-SEP-1998 US 60/100993, 22-APR-1999 US 60/130628 PR  
 23-AUG-1999 US 60/150262  
 PI JOSEPH CHAPPELL, KATHLEEN R MANNA, JOSEPH P NOEL, COURTNEY M PI

STARKS  
 PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N5/10,C12N9/12,  
 PC C12Q1/25,G01N33/68/(C12N9/12,C12R1/91),(C12N9/12,C12R1/19),  
 PC C12N15/00,  
 PC C12N5/00,C12N5/00  
 CC abietadiene synthase  
 FH Key Location/Qualifiers  
 FT CDS (3)..(2606).  
 Location/Qualifiers  
 1..2861  
 /organism="Abies grandis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:46611"

FEATURES  
 source  
 Query Match 48.1%; Score 1078; DB 6; Length 2861;  
 Best Local Similarity 68.8%; Pred. No. 1.5e-309;  
 Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

ORIGIN  
 QY 1 ATGTTTCAGTCCATGGCGACGGTGAAACGAATCCATCTGCAATATGATACAGCTTGGGTG 60  
 Db 378 ATGTTTAGATGTATGGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTGGGTA 437  
 QY 61 GCAGGAATTCCTCAATTCACGGCTCTGGTGCACCCCAATTTCCCAAGCGTCTCAATGG 120  
 Db 438 GCAGGAATTCACAGCAGTGTATGGCTCTGACAAACCCCTCACCTTCTGAGACGGTTGAATGG 497  
 QY 121 ATTCTGAACAATCAACTGCCAGATGCTCGTGGGTGAGGAGTGCAATTTTCTGGCGTAT 180  
 Db 498 ATTCTCAAAATCAGTTGAAAGATGGGTCTTGGGTGAGGATTTCTACTTCTTGGCATAT 557  
 QY 181 GACAGAGTTTAAACACTCTGCGCTGCTCTCACTCTCAAAATATGAAATAGAGGCGAC 240  
 Db 558 GACAGAAATCTGCTACACTTGCATGATTAATTAACCTTACCCTCTGGCGTACTGGGGAG 617  
 QY 241 ATTCAAGTCAGAAAGGGTTGAGTTGTGAGAAACACATGGAAGAAATGAAGAGCAA 300  
 Db 618 ACACAAAGTACAGAAAGGTATTGAATTTCTTCAGGACACAAGCTGGAAGATGAAGTAA 677  
 QY 301 GCTGACAAATCACAGGCCAAAGTGGATTGAGAGTGGTCTGCTGCAATGTTAGATGAAGCA 360  
 Db 678 GCTGATAGTCATAGGCCAAGTGGATTGAAATAGTATTTCTGCAATGCTAAAGGAAGCT 737  
 QY 361 AAAAGCTTGGGATTTGATCTTCTTATCACCTCCCTTTTCACTCTCCAAATFCCAACAAAAG 420  
 Db 738 AAAATCTTAGGCTTGGATCTGCTTACGATTTGGCCATTCCTGAAACAAATCATCGAAAAG 797  
 QY 421 CGCCAGAAAAGCTTCAAAAGATTCCCTCAATGTTTCTCATTAACCATCAGAGCGGCTTG 480  
 Db 798 CGGGAGGCTTAAGCTTAAAGGATTTCCCACTGATGTTCTATATGCCCTTCCAAACAGCTTA 857  
 QY 481 CTCTACTCTCTGAGGGTTTGAAGATGTGGGACTGGCAAGAGATCAAAATCTTCAA 540  
 Db 858 TTGTATTCTTTGGAAGTTTACAGAAATAGTAGACTGGCAGAAATATATGAACCTCAA 917  
 QY 541 TCAAGAGACGATCATTTTAAAGCTCCCTGCACTACTCTGCTTGTGTTCTTCTATGCAACT 600  
 Db 918 TCCAAGGATGGATCAATTTCTCAGCTCTCCGGCATCTACAGCGGCTGTATTTATGCGGTACA 977  
 QY 601 CAAAACAAAGATGCGCTCCACTTTCTCAACTTCTGCTGCTCAGCAAAATTTGGCGGACTAGCTT 660  
 Db 978 GGGAAACAAAAGTGTCTTGGATTTCTTGAACCTTTTGTCTTGAAGAAATTCGGAACCATGTG 1037  
 QY 661 CTTTGCATTTACCCACTTCTATTTTGAACGCTCTGGGCTCTGCGATACAGTTGAGACGC 720  
 Db 1038 CTTTGTCACTATCCGCTTCAATCTATTGTGAACGTTTGTGGCGGTTGATACAGTTGAGCGG 1097  
 QY 721 TTGGGAATCGATCGCTATTTCAGGAAGAAATCAAAAGAAATCTCTGGAATACGTTTATPAGG 780  
 Db 1098 CTAGGTATCGATCGTCAATTTCAAAGAGGAGATCAAGAGGACATTGGATTTATGTTACAGC 1157  
 QY 781 TACTGGGACCGCGAAAGAGCGGTGGGATGGGCAAGATGCAATCCTATTCCTCTGATGTCGAT 840

	D	b	2232	C T C T C A A A A T T G A T C G G A A T C T A G A T T T C T T C A A C T C A T G G C T T T A A C A G G G C G T T T G	2298	
	Q	y	1921	A T C A A T G A C A C A A G A C T T A C C A G C C G A G A G A A A C C G T G G T G N A T T G G T T T C C A G C G T A	1980	
	D	b	2292	G T G A A T G A C A C A A A A C T T A T C A G S C A G A G A G G T C A A G G T G A G G T G G C T T C T G C C A T A	2351	
	Q	y	1981	C A G T G C T A C A T G A G S G A A A T C C G G A G T G C A C A G A G A A A G A A G C T C T A A G T F C A T G T T T A T	2040	
	D	b	2352	C A A T G T T A T A T G A A G B A C C A T C C T P A A A T C T C T G A A G A A G C T C T A C A A C A T G T C T A T	2411	
	Q	y	2041	G G T A T C A T C G A A A C G C A C T G A A G A A T T G A A T T G G G A G T T G G C A C C A C C A G C G A G C A A T	2100	
	D	b	2412	A G T G T C A T G S A A A T G C C T C G A A G A G T T G A A T A G G A G T T G T G A A - - - - - T A A C A A A	2465	
	Q	y	2101	G C C C A T T G T G T G A G A A G A C T G T T C A A C A C T G C A A G A G T G A T G C A G C T G C T G T T T A T	2160	
	D	b	2466	A T A C C G G A T A T T A C A A A A G A C T G G T T T T G A A C T G C A A A T A A A A G A C A T G T C A A A A T	2595	
	Q	y	2161	A T G T A C A G A G T G G C T T T T G T A T C T C - - - T G A C A A A G A G A T G A A A G A C C A T G T C A G C G A	2217	
	D	b	2526	A T G C A A G G G A T G G T T T G A C A C A T A C A T G A T A T G G A A A T T A A A G A G C A T G T C A A A A T	2585	
	Q	y	2218	A C T C T T T T C G A T C C T G T G G C G T A G	2241	
	D	b	2586	T G C C T C T C C A A C C A G T T G C C T A G	2609	
			R E S U L T     3			
			AR240711			
		L O C U S	Sequence 55 from patent US 6468772.	linear	PAT 20-DEC-2000	
		A C C E S S I O N	AR240711			
		V E R S I O N	AR240711.1	GI:27285860		
		K E Y W O R D S	Unknown.			
		S O U R C E	Unclansified.			
		O R G A N I S M	1 (bases 1 to 2861)			
		R E F E R E N C E	Chappell, J., Manna, K.R., Noel, J.P. and Starke, C.M.			
		A U T H O R S	Methods of making modified polypeptides			
		T I T L E	Patent: US 6468772-A 55 22-OCT-2002;			
		J O U R N A L	Location/Qualifiers			
		F E A T U R E S	1..2861	/organism= "unknown"	source	
			/mol_type= "genomic DNA "			
		O R I G I N				
			Query Match	48.1%; Score 1078; DB 6; Length 2861;		
			Best Local Similarity	68.8%; Pred. No. 1-Se-309;		
			Matches 1544; Conservative	0; Mismatches 685; Indels 15; Gaps 4		
	Qy	1	A T G T T T C A G T C C A T G G C G A C G G T G A A A C G A A T C C A T C T C A T A T G A T A C A G C T T G G G T G	437G	60	
	D	b	378	A T G T T T A G A T G T A T G G G C T A T G G C G A A A C G A A T C C C T C T G C A T A T G A C A T G C T T G G G T A	437A	
	Qy	61	G C A A G A T T C G T C A A T T G A G C G C T G T G T G C A C C C A A T T T C C C A A A C G C T T C A A T G G	120T		
	D	b	438	G C A A G A T T C C A G C A G T T G A T G G C T C T G A C A A C C C T C A C T T C T C T G A G A C G G T T G A A T G G	499T	
	Qy	121	A T T C T G A A C A A T C A A C T G C C A G A T G G C T G G G G T G A G A G T G C A T T T T T C T G G C G T A T	180T		
	D	b	498	A T T C T C A A A T C A G T T G A A A G A T G G G T C T T G G G G T G A A G A T T C T A C T T C T T G G C A T A T	557T	
	Qy	181	G A C A G A T T T A A A C A C T C T G C C T G C C T C A C T C T C A A A A T A T G A A T A A G A G G G C A C	241T		
	D	b	558	G A C A A A T A C T G G C T A C A C T T G C A T G T A T T A T A C C T T A C C C T C T G C G T A C T G G G A G	61T	
	Qy	241	A T T C A A G T G C A A A A G G G T T G A G T T T G T G A A A A C A C A T G G A A A A T G A A G A C G A A	300T		
	D	b	618	A C A A A G T A C A G A A A G G T A T T G A A T T C T C A G G A C A C A A G C T G G A A A G A T G G A A T G A A	67T	
	Qy	301	G C T G A A A T C A C A G G C C A A G T G G A T T C A G A G T C G T G T T T C C T G C A A T C T T A G A T G A A G C A	361T		

Db 678 GCTGATAGTCATAGGCCAAGTGGATTGAAATAGTATTTTCTGCAATGCTAAAGGAAGCT 737  
Qy 361 AAAAGCTTGGATGGATCTTCTTATACACCTCCCTTTCATCTCCCAATCCACAAAG 420  
Db 738 AAAATCTTAGCTTGGATCTGCTTACGATTTGCCATTTCTGAAACAAATCATCGAAAG 797  
Qy 421 CGCAGAGAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCTAATACCATCAGACGGGTTG 480  
Db 798 CGGAGGCTAAGCTTAAAGGATTTCCACTGATGTTCTCTATGCCCTTCCMACACGTTA 857  
Qy 481 CTCTACTCTCTGAGGGTTTGAAGATGTTGGAGTGGCAGAGATCAAAATCTTCAA 540  
Db 858 TTGTAATCTTTGGAAGTTTACAGAAATAGTAGACTGGCAGAGAAATTAATGAATCTTCAA 917  
Qy 541 TCAAGAGACGATCATTTTAAAGTCCCTCCCTGATCTACTGCTTGTGTTCTTATGACACT 600  
Db 918 TCAAGAGTGGATCATTTCTCAGCTCTCCGGCATCTACAGCGGCTGTATTTATGCGTACA 977  
Qy 601 CAAACAAAGATGCTCCACTTTCTCACTTCGCTCAGCAAAATTTGCGCACTACGTT 660  
Db 978 GGGMACAAAGTCTGGATTTCTTGAACCTTTGCTTGAAGAAATTCGAAACCATGTG 1037  
Qy 661 CCTTGCATTTACCCACTTGATCTTATTTGAACGCTCTCGGCTCGATACAGTTTGAACGC 720  
Db 1038 CCTTGTCACTATCCGCTTGATCTATTTGAACGTTTGTGGCGGTTGATACAGTTGAGCG 1097  
Qy 721 TTGGGAATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATACGTTTATAGG 780  
Db 1098 CTAGGTATCGATCGTCAATTTCAAGAGGAGATCAAGGAAGCAATGGATTTATGTTACAG 1157  
Qy 781 TACTGGGACCGCAAGAGCGGTGGAGTGGCAAGATGCAATCTTATCTCGATGTCGAT 840  
Db 1158 CATTTGGGA---CGAAAGAGCATTTGATGGCGGAGAGAGAACTCTGTTCTGATATTTGAT 1214  
Qy 841 GACACTGCCATGGGTCTTAGAATCTCTGAGACTTTATGAGATACAAATGATCTTCTCAGATGTT 900  
Db 1215 GATACAGCCATGGGCTTCGAATCTTGAGATTTACATGGATACAAATGATCTTCTCAGATGTT 1274  
Qy 901 CTGGAGAAATTTAGAGACAGAGAAAGAGACTCTTTTGTCTTGGCTTGGCGGTCAACGCAATTT 960  
Db 1275 TTAAGAAACATTTAGAGATGAGAAATGGGAGTTCTTTTGTCTTCTGGGTCAACACAGAGA 1334  
Qy 961 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1020  
Db 1335 GGAATTTACAGATGTTAAACGTCATCGTGTGTTTCAATGTTTCATTTCCGGGAGAAACG 1394  
Qy 1021 AATAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGCCAAAC 1080  
Db 1395 ATCATGGAGAGCAAAACCTCTGTACCGAAAGTATCTGAGGATGCTCTGGNAATGTG 1454  
Qy 1081 AAGCAATTTGATTAAGTGGGCTGTCAAGAGGATCTTCTGGAGAGTGGAGTATGCTATA 1140  
Db 1455 GATGCCCTTTGACAAATGGGCTTTTAAAGAAATATTTCCGGGAGAGGTAGATGTCAC 1514  
Qy 1141 AAGTATCCGTGGATAGAGTATGCCAAGATTTGGAGCAAGATTTACATAGAGCAATTT 1200  
Db 1515 AATAATCCCTGGCATAGAGTATGCCAAGGTTGGAGGCTAGAGCTATATTGGAACACTAT 1574  
Qy 1201 GGATCAAAATGATCTCTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTGACCAAGAA 1260  
Db 1575 GGGCCAGATGATGTGGCTTGGAAACCTGTATATATATATATATATATATGATGCCATCAATTTGCAATGAA 1634  
Qy 1261 AAATATTTGGAGCTGCCAAATTTGACCTTCAATATGTTGCGAGGCTTTACACCAAAAGGAG 1320  
Db 1635 AAGTATTTAGAACTACGCAAACTGGACTTCAATAAGGTGCGTCTATACACCAAAACAGAG 1694  
Qy 1321 ACTCAACATTTGCTAGCTGGTGGAGAGAAATCGGATTTCAATGATCTTACATTCACCCGC 1380  
Db 1695 CTTCAAGATCTTCGAAGGTGGTGGAAATCATCCGTTTTCACGATCTGAATTTCACTCGT 1754  
Qy 1381 CAGCGCCTGTGGAATGATTTCTCAGTGGCGGTTAGTATGTTTTCAGCCAGAAATTCGCT 1440  
Db 1755 GAGCGTGTGACGGAATATATTTCTCACCAGCATCTTTATCTTTGAGCCCGAGTTTCT 1814

Qy 1441 GCTTGTAGAAATTCCTATGCGCAAGACTTCTTGCTCGCAGTATTCTTAGACGATCTTTAC 1500  
Db 1815 AAGTGCAGAGAGGTTTATACAAAACCTTCCAAATTTTCTGTTTATTTTAGATGATCTTTAT 1874  
Qy 1501 GACACCCACGGATCTCTGGATGATCTTAAATTTGTTCTCTGAAGCGGTCCGAAGATGGAT 1560  
Db 1875 GAGCCCATGGATCTTTAGACGATCTTAAGTTGTTTACAGNATCAGTCAAAAGATGGAT 1934  
Qy 1561 ATCTCTGTCTCGATAGCGTTTGGGATTAATCAGTTGAAATTTGCTTCTTAGGGCTGTAC 1620  
Db 1935 CTATCACTAGTGA---CCAAATGCCAACAAATGAAATATGTTTGTGGGTTTCTAC 1991  
Qy 1621 AACACAGTCAATGGATTTGGAAAGATGCACTCAAGGAACNAGGCCGTGATGCTGGGC 1680  
Db 1992 AATACTTTTAATGATATAGCAAAAGAGGACGTGAGAGCAAGGGCGGATGTGCTAGGC 2051  
Qy 1681 TATCTTCGAAAGATAGTGGAGGCGTTCGCTCGCATCGTATACCAAGAAAGCCGAATGCTG 1740  
Db 2052 TACATTCAAATGTTTGGAAAGTCCAACTTCAATGAATACATAGAGATGCGAGTGTCTCA 2111  
Qy 1741 GCAGCAAGATATGTCGCACTTCAACGAATATGTGGAATATGCAAAAGTGTCCATAGCA 1800  
Db 2112 GAAGCTAAATATGTGCCATCTTCAATGAATACATAGAGATGCGAGTGTCTCAATAGCA 2171  
Qy 1801 CTTGCGACAGTGTCTAATAACTCAATCTTTTTCATCTGGAGAAATTAATCTCTGATTAAT 1860  
Db 2172 TTGGAAACAGTGGTCTCATTTAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA 2231  
Qy 1861 TTACAGCAAGTATAGCTTCGGTCCAAATTTCTGCACTTGTCTTGTGCTGCAACGACTA 1920  
Db 2232 CTCTCCAAATTTGATCGCAATCTAGATTTCTTCAACTCATGGGCTTAACAGGGCGTTG 2291  
Qy 1921 ATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
Db 2292 GTGAATGACACCAAACTTTATCAGGCAGAGAGAGGTCAAGGTGAGGTGCTTCTGCCATA 2351  
Qy 1981 CAGTGTCTACATGAGGAAATTCGGAGTGCAAGAGAAAGAGCTCTAAGTCAATGTTTAT 2040  
Db 2352 CAATGTTATATGAAGGACCATCTTAAATCTCTGAAAGAAAGCTCTACAAACATGCTAT 2411  
Qy 2041 GGTATCATCGACCAACGCACTGAAGGAATTTGAATTTGGAGTTGSCCAACCCAGCGACAAT 2100  
Db 2412 AGTGTATGAAATGCTTCCGAGAGTTGAATAGGAGTTTGTGAA-----TAACAA 2465  
Qy 2101 GCCCATTTGTGTGAGAGACTGTGTTTCAACACTGCAAGAGTGAATGCAAGTGTGTTTAT 2160  
Db 2466 ATACCGGATTTTACAAAGACTGGTTTGTGAACTGCAAGAAATATGCAATCTTTTAT 2525  
Qy 2161 ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAAGCAATGTCAAGCGA 2217  
Db 2526 ATGCAAGGGGATGTTTACACTATCAGATGATATGAAATTAAGAGCATGTCAAAAT 2585  
Qy 2218 ACTCTTTTCATCTGTCGGCTAG 2241  
Db 2586 TGCCTCTTCCAAACAGTTGCTAG 2609

RESULT 4  
AR266986

LOCUS AR266986 2861 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 55 from patent US 6495354.

ACCESSION AR266986

VERSION AR266986.1 GI:29696441

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2861)

AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starke,C.M.

TITLE Syntheses

JOURNAL Patent: US 6495354-A 55 17-DEC-2002;

FEATURES Location/Qualifiers

Source

```
1. .2861
/organism="unknown"
/mol type="genomic DNA"
```

## ORIGIN

Query Match	48.1%	Score 1078	DB 6	Length 2861
Best Local Similarity	68.8%	Pred. No. 1.5e-309	Indels 15	Gaps 4
Matches 1544	Conservative	0	Mismatches 685	
Qy	1	ATGTTTTCAGTCCATGGCGACCGGTGAACGAATCCATCTGATATGATACAGCTTGGGTG	60	
Db	378	ATGTTTAGATGATGGCTATGGCAACGAATCCCTCTGCATATGACACTGCTTGGGTA	437	
Qy	61	GCAAGAAATTCGGTCAATTCAGCGCTCTGGTGACCCCAATTTCCCAACGCTTCAATGG	120	
Db	438	GCAAGAAATTCAGCGCTTGGTGACCCCAATTTCCCAACGCTTCAATGG	497	
Qy	121	ATTCTGAACAATCAAATCTGCCAGATGGCTCGTGGGTGAGGAGTGCAATTTTCTGGCGTAT	180	
Db	498	ATTCTTCAAAATCAGTTGAAAGATGGGTCTGGGGTGAAGGATTTCTACTTCTTGGCATAT	557	
Qy	181	GACAGAGTTTAAACACTCTCGCTCGCTTCTCACTCTCAAAATATGAAATGAAGGGCAC	240	
Db	558	GACAGAAATCTGGCTACACTTGCATGATATTTACCCITTAACCTCTGGGTACTGGGAG	617	
Qy	241	ATTCAGTGCAGAAAGGGGTGAGTTTGTGAGAAACACATCGAAGAAATGAAGACGAA	300	
Db	618	ACACAGTACAGAAAGGATTTGAATTTCTTCAGGACACAAGCTGGAAGATGAAGATGAA	677	
Qy	301	GCTGACAAATCAAGGCCAAGTGAATTCGAGGTCTGGTTCCTGCAATGTTAGATCAAGCA	360	
Db	678	GCTGATAGTCATAGGCCAAGTGAATTTGAATAGTATTTCTTGCATGCTTAAGGAGCT	737	
Qy	361	AAAAGCTTGGGATGGATCTTCTTATCACTCCCTTTTCATCTCCAAATCCACAAAG	420	
Db	738	AAAATCTTAGCTTGGATCTGCCCTTAGATTTGCCATTTCTGAAAACAAATCATCGAAAAG	797	
Qy	421	CGCAGAAAGAGCTTCAAAGATTCGCCCTCAATGTTCTTCATAACCATCAGACGGCGTGG	480	
Db	798	CGGAGGCTAAGCTTTAAAGGATTTCCCACTGATGTTCTCTATGCCCTTTCACAACAGTTA	857	
Qy	481	CTCTACTCTCTGGAGGGTTTGCAAGATGTGGTGGACTGGCAGAGATCAAAATTTCAA	540	
Db	858	TTGTATCTTTTGGAGGTTTACAAGAAATAGTAGACTGGCAGAGAAATTAATGAACATTCAA	917	
Qy	541	TCAAGAGCGGATCAATTTTAAAGCTCCCTGCACTACTGCTTGTGCTTCATGACACT	600	
Db	918	TCCAAGGATGATCAATTTCTCAGCTCTCCGGCATCTACAGCGCTGTATTTCATGCGTACA	977	
Qy	601	CAAAAACAAACGATGCTCCACTTTCTCAACTTCGTGCTCAGCAAAATTTGGCGATACGTT	660	
Db	978	GGGACAAAGAGTGCTTGGATTTCTTGAATTTGTCTTGAAGAAATTCGGAAACCAATGTG	1037	
Qy	661	CTTGGCATATCCACTTGTATTTGAACGCTCTGGGCTGTCGATACAGTTGAAACGC	720	
Db	1038	CCTTGTCACTATCCGCTTGTATTTGAACGTTTGTGGGGGTTGATACAGTTGAGCGG	1097	
Qy	721	TTTGGGAATCGATCGCTATTTTCAAGAAAGAAATCAAAGAAATCTCTGGATTAAGTTATAGG	780	
Db	1098	CTAGGTATCGATCGTCAATTTCAAAGGAGATCAAGGAAGCATTTGGATTTATGTTTACAGC	1157	
Qy	781	TACTGGGACGCCGAAGAGCGGTGGATGGGCAAGTGCATCTTATTTCTTGATGTCGAT	840	
Db	1158	CATTGGGA---CGAAGAGGCAATGGATGGGCGAGAGAAATCTGTTCTTGATATTGAT	1214	
Qy	841	GACACTGGCAATGGGCTCTTGAAGATCTTGAGACTTCAATGGATCAATGTATCTTCAGATGTT	900	
Db	1215	GATACAGCCATGGGCCCTTCGATCTTGAGATTTACATGATGATCAATGTATCTTCAGATGTT	1274	
Qy	901	CTGGAGAAATTCAGACGAGAAAGAGACTTCTTTTGTCTTTGCGGTCAAAACGCAAT	960	
Db	1275	TTAAAAACATTTTAGAGATGAGATGGGAGTCTTTTGTCTTTTGGGTCAAAACAGAGA	1334	

Qy	961	GGTGTGCCGATATCTTAACTTTATAGATGTCTTCAAGTATGTCTTCCGGAGAAAAG	1020
Db	1335	GGAGTTACAGACATGTTAAACGTCAATCGTGTTCATCTTCCGGAGAAAACG	1394
Qy	1021	ATAATGGAAGAGCTTAAGACCTTCACTACAAATCATCTCCAAAACTCTTTCGCAAAAAC	1080
Db	1395	ATCATGGAAGAAGCAAAACTCTGTACCGAAAGGTATCTCGAGGAATGCTCTGGAATAATGTG	1454
Qy	1081	AACGCAATTTGATNAGTGGGCTGTCAGNAGGATCTTCTGGAGAGGTGGAAGTATGCTATTA	1140
Db	1455	GATGCTTTTGAACAAATGGCTTTTAAAGAAATATTCGGGAGAGGTAGATATGCACTC	1514
Qy	1141	AAGTATCCGTGGCATAGAAAGTATGCCAAGATTTGGAGGCAAGAGTTTACATAGAGCAATTT	1200
Db	1515	AAATATCCCTGGCAATAAGATATGCCAAGTTTGGAGGCTAGAGCTATATTTGAAACTAT	1574
Qy	1201	GGATCAAAATGATGTCTGGCTGGGAGACGTGTATATAAGATGCTATATGTGAGCAACGAA	1260
Db	1575	GGCCAGATGATGTGTGGCTTGGAATACTGTATATATGATGCAATACATTTCCGATGAA	1634
Qy	1261	AAATATTTGGAGCTGGCCAAAATTTGGAATTTCAATATGTTGAGGCCCTTACACCAAAAGGAG	1320
Db	1635	AGTATTTTAGAACTAGCGAAACTGGACTTCAATAAGGTGCAGTCTATACACCAACACAGAG	1694
Qy	1321	ACTCAACACATTTGACGTGTGGAGAGAAATCGGGATTCGAATCTTAATCTTACACCGC	1380
Db	1695	CTTCAAGATCTTCGAAGTGTGGTGGAAATCATCCGGTTTACGGATCTGAATTTCACTCGT	1754
Qy	1381	CAGCGCCTGTGGAAATGATTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCGCT	1440
Db	1755	GAGCGGTGACGGAATAATATTTCTCACCGGCATCTTTATCTTTGAGCCGAGTTTTCT	1814
Qy	1441	GCCTTGAGAAATTCGCTATGCCAAGACTTTTGCTTCGAGTTATCTTAGACGATCTTTTAC	1500
Db	1815	NAGTCAGAGAGGTTTTATACAAAACCTTCCAAATTTCACTGTATTTTAGATGATCTTTAT	1874
Qy	1501	GACACCCACGGATCTCTGGATGATCTTAATTTGTTCTCTGAAGCGGTCCGAAAGATGGGAT	1560
Db	1875	GACGCCCATGGATCTTTTAGACGATCTTAAGTTGTTTACAGAAATCAGTCAAAAGATGGGAT	1934
Qy	1561	ATCTCTGTGCTGGATAGCTGTTCCGGATAATCAGTTGTAAGATTTGCTTCTTAGGGCTGTAC	1620
Db	1935	CTATCACTAGTGA---CCAAATGCCACAAATAATGAAATATGTTTGTGGGTTTCTAC	1991
Qy	1621	AACACAGTGAATGGAATTTGAAAGATGGACTCAAGGAACAAGGCGGTGATGCTGCTGGGC	1680
Db	1992	AATACTTTTATGATATAGCAAAAGAGGACGTGAGAGCAAGGCGGATGCTGTAGGC	2051
Qy	1681	TATCTTCGAAAGTATGGGAGGGCTTGCTCGCATCGTATACAAAGAGCCGAATGGTCG	1740
Db	2052	TACATTCAAAATGTTTGGAAAGTCCAACTTGAAGTCCAACTTGAAGCTTACAGAAAGACAGAAATGGTCT	2111
Qy	1741	GCAGCAAAAGTATGCGGACATTCACGAATATGTGHAATATGCCAAAGTGTCCATAGCA	1800
Db	2112	GAAGCTAAATATGTGCCATCTCTCAATGAATACATAGAGAATGCCAGTGTGTCAATAGCA	2171
Qy	1801	CTTTCGGACAGTCGTACTTAACTCAATCTTTTCTACTGGAGAAATTTACTTCTCGATTACATT	1860
Db	2172	TTGGAAACAGTCGTTCTCATTTAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA	2231
Qy	1861	TTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCAATCTTTGTGTCTTTTGACTGACGACTA	1920
Db	2232	CTCTCCAAAATTGATCGCGAATCTAGATTTCTTCAACTCATGGGCTTAAACAGGGCGTTTG	2291
Qy	1921	ATCAATGACCAAGACTTTACAGGCCGAGAGAAACCGTGGTGAAATTTGGTTTCCAGCGTA	1980
Db	2292	GTGATGACACCAAACTTATCAGGCAGAGAGGTCAGGTGAGGTGGCTTCTTGCCATA	2351
Qy	1981	CAGTCTCATATGAGGAAAATCCGGAGTGCACAGAGGAAGAAGCTCTTAAGTCTATGTTTAT	2040
Db	2352	CNAUTTTATATGAAGGACCATCTTAAATCTCTGAAGAAAGAAGCTCTTACAAACATGTCTAT	2411
Qy	2041	GGTATCATGCACAAACGCACTGAAGNATTTGAATTTGGAGTTTGGCCAAACCCAGCGAGCAAT	2100



Db 2412 AGTGTATGGAATGCCCTCGAGAGTTGAATAGGAGTTGTGAA-----TAACTAA 2465  
 QY 2101 GCCCAATGTTGTGAGAGACTGCTGTTCAACACATGCAAGAGTGAAGAGCTGTTTAT 2160  
 Db 2466 ATACCGGATATTACAAAGAGCTGGTTTGAAGCTGCAAGATAATGCACTCTTTAT 2525  
 QY 2161 ATGTACAGAGATGGCTTGGTATCTC---TGACAAAGAGATGAAGACCATGTCAGCCGA 2217  
 Db 2526 ATGCAAGGGATGGTTTGACACTATCACATGATGGAATTAAGAGAGCATGCAAAAT 2585  
 QY 2218 ACTCTTTTCGATCCTGTGGCGTAG 2241  
 Db 2586 TGCTCTTCCACCACTGGCTAG 2609

RESULT 5  
 AR316335  
 LOCUS AR316335 2861 bp DNA linear PAT 12-JUN-2003  
 DEFINITION Sequence 55 from patent US 6559297.  
 ACCESSION AR316335  
 VERSION AR316335.1 GI:31711070  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 2861)  
 AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.  
 TITLE Syntheses  
 JOURNAL JOURNAL  
 FEATURES  
 source 1..2861  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Query Match 48.1%; Score 1078; DB 6; Length 2861;  
 Best Local Similarity 68.8%; Pred. No. 1.5e-309;  
 Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;  
 QY 1 ATGTTTCAGTCCATGGCGACCGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTG 60  
 Db 378 ATGTTTAGATGTATGGCTATGGGAAACGAATCCCTCTGCATATGACACTGCTGGGTA 437  
 QY 61 GCAAGAAATTCGTCATAGCGGCTCTGTGACACCCCAATTTCCCAACAGCTTCAATGG 120  
 Db 438 GCAAGGAATTCAGCAGTGTATGGCTCTGCAACCCCTCACTTCTCGAGACGGTTGAATGG 497  
 QY 121 ATTCTGAACAATCACTGCCAGATGGCTCGTGGGGTGAGAGTGCAATTTTCTGGCGTAT 180  
 Db 498 ATTCCTCAAAATCAGTTGAAGATGGGTCTTGGGGTGAAGGATTTCTACTTCTTGGCATAT 557  
 QY 181 GACAGAGTTTAAACACTCTCGCTCGCTCCTCACTCTCAAAATATGGAATAAGGCGGAC 240  
 Db 558 GACAGAAATCTGGCTACACTTGCATGATATATACCTTTACCTCTGGCGTACTGGGAG 617  
 QY 241 ATTCAAGTCAGAAAGGGTTGAGTTGTGAGAAACACATGGAAGAAATGAAGACGAA 300  
 Db 618 ACAAGATGACAGAAAGGATTTGAATTTCTTCAGGACACAAGCTGGAAGATGAAGTAA 677  
 QY 301 GCTGCAATACAGGCAAGTGGATTCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360  
 Db 678 GCTGATAGTCATAGGCAAGTGGATTTGAATAGTATTTCTGCAATGCTTAAAGGAGCT 737  
 QY 361 AAAAGCTGGGATTTGATCTTCTATACCTCCCTTTCATCTCCCAATCCACCAAAAG 420  
 Db 738 AAATCTTAGGCTTGGATCTGCTTACGATTTGGCCATCTGAAACAAATCATCGAAAG 797  
 QY 421 CGCCAGAAAGGCTTCAAAAGATTTCCCTCAATGTTTCTTCAATACCATCAGACGGGTTG 480  
 Db 798 CGGGAGGCTTAAGCTTAAAGGATTTCCCACTGATGTTCTCTATGCGCTTCCCAACACGTTA 857  
 QY 481 CTCTACTCTCTGAGGGTTTGAAGATGTGGTGGAGCTGGAGAGATGACAAATCTTCAA 540

Db 858 TTGTATCTTTGGAGGTTTACAGAAATAGTAGACTGGCAGAAATATATGAACCTTCAA 917  
 QY 541 TCAAGAGACGGATCATTTTTAAAGCTCCCTGCACTTACTGCTTGTGTCTTCTTATGCACT 600  
 Db 918 TCCAGGATGGATCATTTCTCAGCTCTCCGGCATCTACAGCGCTGCTATTCATCGGTACA 977  
 QY 601 CAAACAAAGATGCTTCCACTTCTCAACTTCTGCTCAGCAAAATTTGGCGGATGATGTT 660  
 Db 978 GGGAAACAAAGTGGCTGGATTTCTTGAACCTTTGTCTTGAAGAAATTCGGAACCAATGTG 1037  
 QY 661 CTTTGGCAATTTACCACTTCTGATCTTATTTGAACGCTCTGGCTGCTGATACAGTTGAACGC 720  
 Db 1038 CTTTGTCACTATCCGCTTGAATCTATTTGAACGTTTGTGGCGGTGATACAGTTGACCGG 1097  
 QY 721 TTGGGAATCGATCGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATACGTTTATAGG 780  
 Db 1098 CTAGGTATCGATCGTCAATTTCAAGAGGAGATCAAGAAAGCATTTGGATTTATTTACAGC 1157  
 QY 781 TACTGGGACCGCAAGAGAGCGTGGATGGGCAAGATGCAATCTTATTCCTGATGCGAT 840  
 Db 1158 CATTTGGGA---CGAAAGAGGCAATGGATGGGCGAGAGAAATCTCTGCTGATATTTGAT 1214  
 QY 841 GACACTGCCATGGCTCTTAGAATCTTAGAATCTTAGAATCTTAGAATCTTAGAATCTTAGAAT 900  
 Db 1215 GATACAGCAATGGGCTTCCGAATCTTAGAATCTTAGAATCTTAGAATCTTAGAATCTTAGAAT 1274  
 QY 901 CTGGAGAAATTTCAAGAGACGAGAAAGAGACTTCTTCTTGTCTTGTCCCGGTCAAAACGCAAT 960  
 Db 1275 TTAATAAACATTTAGAGATGAGAAATGGGAGTTCTTTTGTCTTCTGGGTCAACACAGAGA 1334  
 QY 961 GGTGTGACCGATAATCTTAACTTTATAGATGTTTCAAGATGTTTCCGGGAGAGAAAG 1020  
 Db 1335 GGAGTTTACAGACATGTTAAACGTCATCGTTGTTCACTGTTTCTTCTGGGAGAAACG 1394  
 QY 1021 ATAAATGGAAGAGCTAAGACCTTCACTACAAATCACTCCAAATGCTCTTGGCCAAAC 1080  
 Db 1395 ATCATGGAAGAGAAACAACTGTACCGAAAGGATCTGAGGAATGCTCTGGAATAATGTG 1454  
 QY 1081 AACGCAATTTGATTAAGTGGCTCTCAAGAGGATCTTCTCGAGAGAGTGGAGTATGCTATA 1140  
 Db 1455 GATGCTTTGACAAATGGGCTTTTAAAGAAATATTCGGGGAGAGGTAGAGTATGCACTC 1514  
 QY 1141 AAGTATCCGTGGCATAGAAGTATGCAAGATTTGGAGGCAAGATTTCAATAGAGCAATTT 1200  
 Db 1515 AAATATCCCTGGCATAAGAGTATGCCAGGTTGGAGGCTAGAAAGCTATATTTGAAACATAT 1574  
 QY 1201 GGATCAAAATGATGCTGCTGGCTGGGAGAGCTGTGATAAGATGCTATATCTGACCAACGAA 1260  
 Db 1575 GGGCCAGATGATGTGGCTTGGAAAACTGTATATATGATGCCATACATTTTCGAATGAA 1634  
 QY 1261 AAATATTTGGAGCTGGCCAAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGAATTTGGA 1320  
 Db 1635 AAGTATTTAGAACTAGCGAACTGGACTTCAATAGGTGCACTATACACCAACAGAG 1694  
 QY 1321 ACTCAACACATTTGCTGCTGGTGGAGAGAAATCGGATTCATGATCAATTCATTCACCCCG 1380  
 Db 1695 CTTCAAGATCTTCAAGAGTGGTGAATCATCCGTTTCAAGGATCTGAATTTCACTCGT 1754  
 QY 1381 CAGCGGCTGTGGAATGTTATTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCGCT 1440  
 Db 1755 GAGCGTGTGACGGAATATATTTCTACCGGCACTCTTATCTTTTGGAGCCGAGTTTCT 1814  
 QY 1441 GCTTGTAGAAATTTGCTTATCGCAAGACTTCTTGGCTCGCAAGTATTTCTAGACGATCTTTAC 1500  
 Db 1815 AAGTGCAGAGAGTTTATACAAAACTTCAATTTCACTGTTATTTTAGATGATCTTTAT 1874  
 QY 1501 GACACCCAGGATCTCTGATGATCTTAAATGTTCTCTGAAAGCGGTCCGAAAGATCGGAT 1560  
 Db 1875 GACGCCCATGGATCTTTAGACGATCTTAAGTTGTTTACAGAAATCACTGCAAAAGATGGAT 1934  
 QY 1561 ATCTGTGCTGATAGCGTTCCGGGATTAATCAGTTGGAAGTTGCTTCTCTAGGCTGTAC 1620

Db	1935	CTATCACTAGTGA---CCAAATGCCAACAAATGAAAATATGTTTTGTGGGTTTCTAC	1991
Qy	1621	AACACAGTGAATGGATTTGGAAAGATGACCTCAAGGAACAAGCGCGTGTATGTCTGGGC	1680
Db	1992	AATACTTTTAATGATATAGCAAAAGAGACGTGAGAGGCAAGGGCGGATGTCTAGGC	2051
Qy	1681	TATCTTCGAAAGATGATGGAGGCTTGTCTGCATCGTATACCAAGAGCGGAATGGTGC	1740
Db	2052	TACATTCAAATGTTTGGAAAGTCCAACTTGAAGCTTACAGAAAGAGCAGAATGGTCT	2111
Qy	1741	GCAGCAAGATGATGGCGACATTCACAGAAATATGTGAAATATGCCAAAGTTCATAGCA	1800
Db	2112	GAAGCTAAATATGTGCGCATCTTCAATGAATATACATAGAGATGCGAGTGTCAATAGCA	2171
Qy	1801	CTTGGCAGCTGCTACTAACTCAATCTTTTCACTGGAGAAATTAATCTCTGATTAAT	1860
Db	2172	TTGGGAACAGTCGTCTCATTAATAGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA	2231
Qy	1861	TTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCATCTTGTGCTTTTGTAGCTGAGCACTA	1920
Db	2232	CTCTCCAAATTGATCGGAATCTAGATTTCTTCAACTCATGGGCTTAAACAGGGCGTTTG	2291
Qy	1921	ATCAATGACCAAGACTTACCAGCGGAGAGAACCGTGTGAATTTTCCAGCGTA	1980
Db	2292	GTGAATGACACCAAACTTATCAGGCAGAGAGAGGTCAAGGTGAGTGGCTTCTGCCATA	2351
Qy	1981	CAGTGTACATGAGGGAATCCGGAGTGCAAGAGGAAGAGCTCTAAGTCATGTTTAT	2040
Db	2352	CAATGTTATATGAAGGACCATCTTAATCTCTGAAGAGAGGCTCTACAACTGTCTAT	2411
Qy	2041	GGTATCATCGCAACGCACTGAAGAAATGAAATTTGGGAGTTGGCCAAACCCAGCGCAAT	2100
Db	2412	AGTGTATGGAATATCCCTCGGAAGTTGAATAGGGAGTTTGTGA-----TAACAA	2455
Qy	2101	CCCCATTTGTGTGAGAGAGCTGCTTTCAACACTGCAAGAGTATGATGAGCTGTTTAT	2160
Db	2466	ATACCGGATATTTACAAAGACTGGTTTTTGGAACTGCAAGATATGCAACTCTTTAT	2525
Qy	2161	ATGTACAGAGATGGCTTTGGTATCTC---TGCAAAAGAGATGAAGACCATGTGAGCGGA	2217
Db	2526	ATGCAAGGGATGTTTGCACATATCACATGATATGGAATTAAGAGCATGTCTCAAAAT	2585
Qy	2218	ACTCTTTTGCATCTGTGGCTAG 2241	
Db	2586	TGCCTCTTCCACCAAGTTGCCTAG 2609	
RESULT 6			
AR338478			
LOCUS 2861 bp DNA linear PAT 17-AUG-2003			
DEFINITION Sequence 55 from patent US 6569656.			
ACCESSION AR338478			
VERSION AR338478.1 GI:33725255			
KEYWORDS Unknown.			
SOURCE Unknown.			
ORGANISM Unclassified.			
REFERENCE 1 (bases 1 to 2861)			
AUTHORS Chappell, J., Manna, K.R., Noel, J.P. and Starke, C.M.			
TITLE Syntheses			
JOURNAL Patent: US 6569656-A 55 27-MAY-2003;			
FEATURES			
source Location/Qualifiers			
1..2861			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Query Match 48.1%; Score 1078; DB 6; Length 2861;			
Best Local Similarity 68.8%; Pred. No. 1.5e-309;			
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;			
Qy	1	ATGTTTCAGTCCATGGCGAGCGGTGAACCAATCCATCTGCATATGATACAGCTTGGGTG	60



QY	661	CCTTGCCATTACCCACTTGATCTATTGTAACGCGCTCTGGGCTGTGATACAGTTGAACGC	720
Db	1038	CTTGTGCTATCCGCTGATCTATTGTAACGCTTGTGGCGGTTGATACAGTTGAACGC	1097
QY	721	TTGGGAATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGATTAAGTTATAGG	780
Db	1098	CTAGGTATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGATTAAGTTATAGG	1157
QY	781	TACTGGGACGCCAAGAGCGTGGGATGGGCAAGATGCAATCTATTCTCTGATGCGAT	840
Db	1158	CATTGGGA---CGAAGAGGCAATGGATGGGCGAGAGAGATCTGTCTCTGATATTGAT	1214
QY	841	GACATGCCATGGGTCTTAGAATCCTGAGACTTCTCATGGATACAAATGATCTCTCAGATGTT	900
Db	1215	GATACAGCCATGGGCTCTGAAATCTTGAGATTACATGGATACAAATGATCTCTCAGATGTT	1274
QY	901	CTGGAGAAATTCAGACGAGAAAGGAGACTCTCTTTGCTTTGCGGTCGCGTCAAGCGCAAAAT	960
Db	1275	TTAAAGAAATTCAGACGAGAAAGGAGACTCTCTTTGCTTTGCGGTCGCGTCAAGCGCAAAAT	1334
QY	961	GGTGTGACCGGATATCTTAACCTTTATAGATGTTCAAGATGTTCTTCCGGGAGAAAG	1020
Db	1335	GGAGTTACAGACATGTTTAAAGCTCAATCGTGTGTTTCAATGTTTCAATTTCCGGGAGAAAG	1394
QY	1021	ATATGGAAGAGTAAGACTTCACTACAAATCTCTCAAAATGCTCTGCGCAAAAC	1080
Db	1395	ATCATGGAAGAGTAAGACTTCACTACAAATCTCTCAAAATGCTCTGCGCAAAAC	1454
QY	1081	AACGCATTTGATAGTGGGCTCTCAAGAAAGATCTCTCGAGAGGTTGGATATGCTGATA	1140
Db	1455	GATGCTTTGCAAAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGATATGCACTC	1514
QY	1141	AAGTATCCGTGGCATAGAAATGCAAGATGTCGAAGATGTCGAAGATGTCGAAGATGTC	1200
Db	1515	AAATATCCGTGGCATAGAAATGCAAGATGTCGAAGATGTCGAAGATGTCGAAGATGTC	1574
QY	1201	GGATCAATGATGCTGCTGGGAGAGACTGTATAGATGCTATATGTCGAGCAACGAA	1260
Db	1575	GGGCGAGATGATGCTGGGCTTGGAAATCTGTATATATATGATGATGATGATGATGAT	1634
QY	1261	AAATATTTGGAGCTGGCCAAATGGACCTTCAATATGTCGAGGCTTATACAAAGAG	1320
Db	1635	AAGTATTTAGAACTAGCGAAATCGACTTCAATAGGTGCGATCTATACCAACACAG	1694
QY	1321	ACTCAACATGTCAGTGTGGAGAGAAATCGGATTCATGATCTTACATTTACCCGC	1380
Db	1695	CTTCAAGATCTTCAAGGTGGTGGAAATCATCCGGTTTACCGATCTGAAATTTCACTGT	1754
QY	1381	CAGCGGCTGTGGAATGATTTCTCAGTGGCGGTTAGTATGTTTGGAGCAGAAATTCGCT	1440
Db	1755	GAGCGTGTGCGGAATATATTTCTCACCGGCATCTTTATCTTTGAGCCGAGTTTCT	1814
QY	1441	GCTTGTAGAAATTCCTATGCCAAGACTTCTGCGTGGCGGTTAGTATGTTTGGAGCAGAAATTCGCT	1500
Db	1815	AAGTGCAGAGAGTTTATACAAAAATCTCAATTTTCACTGTTATTTTAGATGATCTTTAT	1874
QY	1501	GACACCCAGGATCTCTGATGATCTTAAATTTGTTCTTGAAGCGGTCCGAAGATGGGAT	1560
Db	1875	GAGCGGCTGGATCTTTAGCAGATCTTAAATTTGTTTCAAGATCAGTCAAAAGATGGGAT	1934
QY	1561	ATCTCTGTCTGATAGCGTCTGGGATATCAAGTGAAGTTGCTTCTAGGCTGTATC	1620
Db	1935	CTATCACTAGTGA---CCAAATGCCAACAAATGAAATATGTTTGTGGGTTTCTAC	1991
QY	1621	AACACAGTGAATGGATTTGGAAGATGATGACTCAAGGAACAGGCGGTGATGCTGGGC	1680
Db	1992	AATACTTTTAAATGATATAGCAAAAGAGGACGTGAGAGGCAAGGCGGATGCTAGGC	2051
QY	1681	TATCTTTCGAAAGATGATGGAGGCTTGTCTGCGATCTGATATCAAGAAAGCCGAATGGTCTG	1740
Db	2052	TACATTCAAATGTTTGGAAAGTCCAACTTGAAGCTTACAGGAAGAGCAGAAATGGTCT	2111
QY	1741	GCAGCAAGATGATGCGGACATTCAGAGAAATATGTGGAATAATGCCAAAGTGTCTCATGCA	1800

RESULT 8  
 AGU50768 2861 bp mRNA linear PLN 02-AUG-1996  
 LOCUS Abies grandis abietadiene synthase (ac22) mRNA, complete cds.  
 DEFINITION  
 ACCESSION U50768  
 VERSION U50768.1 GI:1477583  
 KEYWORDS  
 SOURCE Abies grandis  
 ORGANISM Abies grandis  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 REFERENCE 1 (bases 1 to 2861)  
 AUTHORS Stoffer-Vogel, B., Wildung, M.R., Vogel, G. and Croteau, R.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (06-MAR-1996) Mark Wildung, IBC, Washington State  
 University, Pullman, WA 99164-6340, USA  
 FEATURES  
 Location/Qualifiers  
 1..2861  
 /organism="Abies grandis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:46611"  
 /db\_xref="GI:1477583"  
 /tissue\_type="wounded stem"  
 1..2861  
 /gene="ac22"  
 3..2609  
 /gene="ac22"  
 /note="diterpene cyclase"  
 /product="abietadiene cyclase"  
 /protein\_id="AAB05407.1"  
 /translation="MAMPSSLSLSQIPFAAHLTANAQSIPIHPTTLNAGSSASKRRS  
 LVLWKGSKNIIACVGGGATSVQSAEKNDLSSTLVKREFFPPGWKDLIDSL  
 TSHKVAASDEKRIETLISEIKNMPGCVGETNPAYDTAWTAVIPAVDGSNPHFP  
 ETEWLNQLKDGSGWGEFYLAVDRILATLACITLTLTWRTGETQVKGIFBFRQ  
 AGMDEADSHRPSGFEIVFPAMLEKALIGLDLPDLPELKQIIEKREAKLRIPD  
 VLVALPTLLYSLGLEQIIVDWQIMKIQSKGDSFLSSPASTAAVPMRTGNKKCLDFL



Db	2292	GTGAATGACACCAAACTTATCAGGCAGAGAGGTCAAGGTGAGTGGCTTCTGCCATA	2351	Db	194	CGCCCGCAATTTCCCAACAGTTGACTGGATTTTGAATAACCCAGTTAAAGATGGTTCA	253
Qy	1981	CAGTGTCTACATGAGGGAATTCGGAGTGCACAGAGAGAAAGCTCTAAGTCAATGTTTAT	2040	Qy	151	TGGGTGAGGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTCGCTC	210
Db	2352	CAATGTTATAGAGACCATCTAAATCTCTGAAGAGAGCTCTCAACATGTCAT	2411	Db	254	TGGGGAATTCAGTCCACCTTCTGCTGTCGACCGTCTTCTGGCACTCTTCTTGTGTT	313
Qy	2041	GGTATCATCGAACACGACATGAAGGAATTAATTTGGAGTTGGCCAAACCCAGCGCAAT	2100	Qy	211	CTCACTCTCAAAATATGGAATTAAGGCGCAATTCAGAGTCAGAAAGGGTGTAGTTGTG	270
Db	2412	AGTGTATGGAATGCGCTCGAAGAGTTGAATAGGAGTTTGTGA	2465	Db	314	CTTGTGCTCTTAAATGGAACGTTGGGATCTCAAGTAGAGAGGGAATTTGAATTCATA	373
Qy	2101	GCCCCATTGTGTGAGAGACCTGTTTCAACACTGCAAGAGTGAAGCACTGTCTTTAT	2160	Qy	271	AGAAAACATGGAAGAAATGAAGGACGAGCTGACATCACAG--GCCAAGTGAATC	327
Db	2466	ATACCGGATATTTACAAAGACCTGTTTGAACCTGCAAGATATGCAACTCTTTAT	2525	Db	374	AGAGCAATCTGGAACCTAGTAAAGGATGAACCGATCAAGATAGCTTGGTAAACAGACTTT	433
Qy	2161	ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCAATGTCCAGCGA	2217	Qy	328	GAGTGTGTTTCTCTCAATGTTTAGATGAAGCAAAAGCTTTGGATTTGATCTTCTTAT	387
Db	2526	ATCAAGGGGATGGTTTGACACTATCATGATATGGAATTAAGAGCATGTCAAAAT	2585	Db	434	GAGATCAATATTTCTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGACTTCCCTAC	493
Qy	2218	ACTCTTTTCATCTGTGGCGTAG	2241	Qy	388	CACCTCCCTTTCATCTCCCAATCCCAAGGGCCAGAAAGCTTCAAAAGATCCCC	447
Db	2586	TGCTCTTCCAACAGTTGCTAG	2609	Db	494	GACCTGCTTATATACATCTGTTGCAGACTAAACGGCAGGAAGATTAGCAAACTTTCA	553
RESULT 9				Qy	448	CTCAATGTTCTTCAATACCATCAGACGGCTGCTCTACTCTCTGGAGGGTTTGAAGAT	507
LOCUS	BD227671	2424 bp	DNA	linear	PAT 17-JUL-2003		
DEFINITION	BD227671						
ACCESSION	BD227671						
VERSION	BD227671.1	GI:33037441					
KEYWORDS	JP 2002526066-A/23.						
SOURCE	Abies grandis						
ORGANISM	Abies grandis						
REFERENCE	1	(bases 1 to 2424)					
AUTHORS	Chappell, J., Manna, K.R., Noel, J.P. and Starks, C.M.						
TITLE	Syntheses						
JOURNAL	Patent: JP 2002526066-A 23 20-AUG-2002;						
COMMENT	UNIVERSITY OF KENTUCKY RESEARCH DEPARTMENT, THE SALK INSTITUTE FOR BIOLOGICAL STUDIES SECRETARY DEPARTMENT OF HEALTH AND HUMAN SERVICES						
	OS Abies grandis (giant fir)						
	PN JP 2002526066-A/23						
	PD 20-AUG-2002						
	PF 17-SEP-1999 JP 2000574228						
	PR 18-SEP-1998 US 60/100993, 22-APR-1999 US 60/130628						
	PI JOSEPH CHAPPELL, KATHLEEN R MANNA, JOSEPH P NOEL, COURTNEY M STARKS						
	PC C12N15/09, A01H5/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N5/10, C12N9/12						
	PC C12Q1/25, G01N33/68// (C12N9/12, C12R1/91), (C12N9/12, C12R1/19),						
	PC C12N15/00,						
	PC C12N5/00, C12N5/00						
	CC E-alpha-bisabolene synthase						
	PH Key Location/Qualifiers						
	FT CDS (2). (2347).						
FEATURES	Location/Qualifiers						
source	1. .2424						
	/organism="Abies grandis"						
	/mol_type="genomic DNA"						
	/db_xref="taxon:46611"						
ORIGIN							
	Query Match 32.3%; Score 723.4; DB 6; Length 2424;						
	Best Local Similarity 59.8%; Pred. No. 8.7e-204;						
	Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;						
Qy	31	AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATCCCGTCAATTTGACGGCTCTGTT	90	Qy	1108	AAGGATCTTCTGGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAGATAGCCA	1167
Db	134	ACTCCATCTGCTTATGACACAGCATGGTAGCGAGGTGCCGCCCATTTGATGCTCTGCT	193	Db	1211	CAAAACCTGAGCAAGAGATCAATACGGCTGAAGACTTCTTGGCATGCCAGTGTCCG	1270
Qy	91	GCACCCCAATTTCCCAACAGCTTCAATGATTTCTGNACAAATCACTGCCAGATGGCTCG	150	Qy	1168	AGATTGGAGGCAAGAGTTACATAGACCAATTTGGATCAATGATGTCTGGCTGGGAAG	1227

Db 1271 AGAGTGGNAGCAAGAGATACTGTCAAGTGTATCGCCAGATTATGACGCATAGCAAAA 1330  
Qy 1228 ACTGTGTATGAAGTCTATATGTAGCAACGAAAAATATTTGGAGCTGGCCAAATTTGGAC 1287  
Db 1331 TGGCTTTTACAAAGTCACTCCGTGAAATGAAGAGTTTTTAGAGCTGGGAAAAATTAGAT 1390  
Qy 1288 TTCAATATGTGTGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTGAGCTGGTGAGA 1347  
Db 1391 TTCAACATTTACAGTCCATCCACCAAGAAAGAAATGAAGATGTTTACCAGCTGGTTTGA 1450  
Qy 1348 GAATCGGATTTCAATGATCTTACATTCACCCGCGAGCTGTGGAAATGATTTCTTCA 1407  
Db 1451 GATTCGGGTGGCCACTTTCACCTTCGCTCGGGAGAGCGCTGGAATTTCTACTTCTTA 1510  
Qy 1408 GTGCGGTGTATGTTGTGAGCAGAAATTCGCTGTGTGTAGAAATGCTGTATGCCAAGACT 1467  
Db 1511 GTAGCGCGGGACCTATGAACCCAGATATGCCAAATGCGAGGTTCCTCTTTACAAAAGTG 1570  
Qy 1468 TCTTGCTCGCAGTATTCTAGACGATCTTTACGACACCCACGATCTCTGGATGATCTT 1527  
Db 1571 GCATGCTTGAGACTGTTCGTGAGCAGATATGTATGACACTTTATGGAACCCCTAGATGAATTG 1630  
Qy 1528 AAATTTGCTCTGAAGCGGTCCGAAGATGGATATCTCTGTGTGTGATAGCGTTTCGGAT 1587  
Db 1631 AAGCTATTCACTGAGCTGTGAGAGATGGACCTCTCTTTACAGAAACCTTCCAGAC 1690  
Qy 1588 AATCAGTTGAAATTTGCTTCTTAGGGCTGTACAACACAGTGAATGATTTGGAAGAAT 1647  
Db 1691 TAT---ATGAACTATGTATTACCAATCTATTATGACATAGTTTCCAGAGGTGGCTTGGAG 1747  
Qy 1648 GGACTCAAGAAACAGCGGTGTGTGCTGGGTATCTTCGAAAGATGTGGGAGGCTTG 1707  
Db 1748 GCAGAGAGAAACAGGGCGGTGAATTTGGTCAGCTTTTTCAGAAAGGATGGAGGATAT 1807  
Qy 1708 CTGCTCGTATACCAAGACCGGAATGCTCGGACGAAAGTATGTGCGCAATTTCAAC 1767  
Db 1808 CTCTCGGTATTATGAGAGCTGTAATGTTAGCTGTGAGTATGTGCTACTCTTGGAC 1867  
Qy 1768 GAATATGTGAAATGCAAGTGTCTCATAGCACTTGCGCAGTGTCTATCAACTCAATC 1827  
Db 1868 GAGTACATAAAGAAATGGAATCACAATCTATCGGCAACGATATCTCTGTTGAGTGAGTG 1927  
Qy 1828 TTTTTCAC---TGGAGAAATTTCTCTGATTACATTTTACAGCAAGTATGACCTTCGGTCC 1884  
Db 1928 TTGATAATGGAATGGGCAACTCTTTTCGCAAGAGGCAATTAGAGAAAGTAGAATTTACGGA 1987  
Qy 1885 AAATTTCTGC-----ATCTTGTGTCTTTGACTGGACGACTAATCAATGACACCAAGACT 1938  
Db 1988 AGACGTGTTCTACAGAGCTGAATAGCTCATTTCCGGCTGGCGGATGACACGAAGACA 2047  
Qy 1939 TACCAGCCGAGAGAAACCGTGTGAAATTTGGTTTCCAGCGTACAGTGTACATGAGGGAA 1998  
Db 2048 TATAAAGCTGAGAAGGCTCGTGAGAAATTTGGCTCCAGCATTTGAATGATGAAGAGC 2107  
Qy 1999 AATCCGAGTGACAGAGNAGAGCTCTAAGTCAATGTTATGTTATGATATCGACAACGCA 2058  
Db 2108 CATCTCTGAATGTACAGAGGAAGAGGCTCTCGATCACAATCTATAGCAATTCGAGCGCGG 2167  
Qy 2059 CTGAAGGAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2118  
Db 2168 GTGAGGAAGACTGACAGAGATTTCTGAGCCCG---ACGAGCTCCCATTCGCTCCGAG 2224  
Qy 2119 AGACTGCTGTTCAACTGCAAGAGTGTGACAGTGTGTTTATATGATACAGAGATGCTTTT 2178  
Db 2225 AAGATGCTTTTCAGGAGACAAGAGTGACGATGGTGATATTCAAGGATGAGATGATTC 2284  
Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGACGCGCACTTTTTCGATCTCTGTGGC 2237  
Db 2285 GGTGTTTCCAAATTTGAAGTCAAAGATCATATCAAGAGAGTGTCTCTTGAACCGCTGCC 2343

RESULT 10  
AR222107

LOCUS AR222107 2424 bp DNA linear PAT 26-SEP-2002  
DEFINITION Sequence 15 from patent US 6429014.  
ACCESSION AR222107  
VERSION AR222107.1 GI:23329481  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2424)  
AUTHORS Steele,C.L., Bohlmann,J. and Croteau,R.B.  
TITLE Monoterpene synthases from grand fir (*Abies grandis*)  
JOURNAL Patent: US 6429014-A 15 06-AUG-2002;  
FEATURES Location/Qualifiers  
source 1..2424  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN  
Query Match 32.3%; Score 723.4; DB 6; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 8.7e-204;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;  
Qy 31 AATCCATCTGCATATATACAGCTTGGGTGGCAAGAAATCCCGTCAATTGACGGCTCTGGT 90  
Db 134 ACTCCATCTGCTTATGACACGATGGGTAGCGGGTGGCCGCCATTGATGCTCTGCT 193  
Qy 91 GCACCCCAATTTCCCAACCGCTTCAATGGAATCTCAACAATCAACTGCCAGATGGCTCG 150  
Db 194 CGCCCGCAATTTCCCAACAGTTGACTGGATTTTGAATAAACCCAGTTAAAGATGGTTCA 253  
Qy 151 TGGGGTGAGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCCTGCCTC 210  
Db 254 TGGGGAAATTCAGTCCCACTTTTCTGCTGTCGACCGCTCTTCTTGCCTCACTCTTCTTGTGTT 313  
Qy 211 CTCACCTCTCAAAATATGGAATAAGGGCGCACTTCAAGTGCAGAAAGGGTGTGAGTTGTG 270  
Db 314 CTGTGCTCTTAATGGAACGTTGGGGATCTGCAAGTAGACGAGGAAATGAAATCATTA 373  
Qy 271 AGAAAAACATGGAAGAAATGAAGACGAAGTGTCAACATCAACACAG---GCCAAGTGAATC 327  
Db 374 AAGAGCAATCTGCACTAGTAAAGATGAAACCGATCAAGATAGCTTGGTTAACAAGACTTT 433  
Qy 328 GAGTGTGTTTCTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATGGAATCTTCTCTAT 387  
Db 434 GAGATCATATTTCTCTCTCTGTTAAGAGAAGCTCAATCTCTGCGCTCGGACTTCCCTAC 493  
Qy 388 CACCTCCCTTCAATCTCCCAATCCACCAAAAGCGCCAGAAAAGCTTCAAAAGATTTCCC 447  
Db 494 GACCTGCTTATATACATCTGTTGCAGACTAAACGGCAGGAAGATAGCAAACTTTCA 553  
Qy 448 CTCATGTTTCTTCAACCATCAGACGGGTTGCTCTACTCTCTGGAGGGTTTGCAGAT 507  
Db 554 AGGGAGAAATTTATGCGGTTCGTCGCCATTGTTGTTATTTTATAGAGGGAATACAAGAT 613  
Qy 508 GTGGTGACTGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTAAAGTCC 567  
Db 614 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAGTCAGGATGGGTCTTTTCTTAAGTCA 673  
Qy 568 CTGCACTCTACTGCTGTGCTTCTATGACACTCAAAACAAACGATGCGCTCCACTTCTC 627  
Db 674 CTGCTTCTACTGCTGCTGCTTTTTCATGCAACAGGAGACGGAAATTCCTTGAATCTTG 733  
Qy 628 AACTTCGTCTCAGCAAAATTTGGCGACTAGCTTCTTCTTGGCAATPACCCACTTGATCTT 687  
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGTATCTCTGTGGATCTGCTG 793  
Qy 688 GAACGCTCTGGGCTGTCCATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAAA 747  
Db 794 GAACGCTCTGTATCGTAGATAATATTGTACGCTTTGGAATCTATAGACACTTTGAAAG 853  
Qy 748 GAAATCAAGAAATCTCTGATTTACGTTTATAGTACTGGGACCGCCGAAGAGCGCTGGGA 807  
Db 854 GAAATCAAGAAAGCTCTTGATTTATGTTTACAGGCATTGGAA---CGAAAGAGAAATTTGG 910







QY 388 CACCTCCCTTTCATCTCCCAATCCACCAAGCGCCAGAAAAGCTTCAAAGATTCCC 447  
Db 494 GACCTGCCCTTATATACATCTGTTGCAGACTAAACCGCAGGAAGATTTAGCAAAACTTTCA 553  
QY 448 CTCAATGTTCTTCATAACCATCAGACGGGTTGCTCTACTCTCTCGAGGGTTTGGCAAGAT 507  
Db 554 AGGGAGGAATTTATCGGGTTCGTCGCCATTTGTTGTATCTTTAGAGGGAATACAGAT 613  
QY 508 GTGGTGAGCTGGCAAGAGATACAAATCTTCAATCAAGAGACGGATCATTTTTTAAAGCTCC 567  
Db 614 ATAGTTGAATGGCAAGCAATTAAGAAAGTCAAAAGTCAGAGTGGGTCTTTCTTAAGCTCA 673  
QY 568 CTTGCACTTACTGCTTGTCTTCATGCACTCAAAACAAAGAGTGCCTCCACTTTCTC 627  
Db 674 CTTGCTTCTACTGCTGCGTCTTTCATGCAACAGGAGACGCGAAATGCGCTTGAATTTCTTG 733  
QY 628 AACTTCGTGCTCAGCAAAATTTGGCGACTAGCTTTCCTTGGCATATACCACCTTGATCTATTT 687  
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTCCCTGCTGTATCCTGTGATCTGCTG 793  
QY 688 GAAACGCTCTGGGCTGTCATACAGTTGAACGCTTGGGAATCGATCGCTATTTTCAAGAAA 747  
Db 794 GAAACGCTCTGTGATCGTAGATAATATTGACGCTTTGGAATCTATAGACACTTTTGAAG 853  
QY 748 GAAATCAAGAAATCTCTGATACGTTTATAGTTTATAGTACTGGGACGCGAAAGAGCGTGGGA 807  
Db 854 GAAATCAAGAAAGCTCTGATTTATGTTTACGCAATTGGAA---CGAAAAGAGAAATTGGG 910  
QY 808 TGGGCAAGATGCAATCCTATCTCTGATGTCGATGACACTGCCATGGGTCTTAGAATCCCTG 867  
Db 911 TGGGCGAGACTAAATCCCATAGCAATCTTTGAGACCACTGCTTTGGGATTTGATTTGCTT 970  
QY 868 AGACTTCATGGATACAAATGTATCTTCAGATGTTTCGGAGAAATTCAGAGACGAGAAAGGA 927  
Db 971 CGGCTGATAGGTACAAATGATCTCCAGCAATTTTGAACACTTCAAAGATGCCATGGG 1030  
QY 928 GACTCTTTTGTCTTCCCGTCAACGCAAAATTTGTTGTCACCGATTAATCTTAACTTTAT 987  
Db 1031 AAATTCATTTGCTCGACCGCTCAATTTCAAACAAAGATGTAGCAAGCATGCTGATCTTTAT 1090  
QY 988 AGATGTTCAAGATGATTTTTCGGGAGAAAGATAATGGAAGACTTAAGACCTTCACT 1047  
Db 1091 AGAGCTTCCAGCTCGCATTTCCGGAGAAACATTTCTGTAGAACATAAAGCTTCGCT 1150  
QY 1048 ACAAAATCATCTCAAAATGCTCTTGCAAAAACAAACGCAATTTGATAAGTGGGTGTCAAG 1107  
Db 1151 ACTAAATATTTGAGAGAAAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAAACAA 1210  
QY 1108 AAGGATCTTCTGGAGAGGTGAGTATGCTATAAAGTATCCGTTGGCATAGAGATGCGCA 1167  
Db 1211 CAAAACCTGAGCCACAGAGATCAAAATACGGGCTGAAGACTTCTTGGCATGCCAGTGTCCG 1270  
QY 1168 AGATTTGGAGCGAAGATTACATAGACAAATTTGGATCAATGATGCTGCGCTGGGGAAG 1227  
Db 1271 AGAGTGAAGCAAGAGACTGTCAAGTGTATCGCCAGATTAATGACGCAATAGCAAAA 1330  
QY 1228 ACTGTGTATAAGATGCTATATGTAGCAACGAAAAATATTTGGAGCTGGCCAAATTTGGAC 1287  
Db 1331 TGGGTTTACAGTACCTACGTGAACATGAAAGCTTTTAGAGCTGGGAAAAATTAGAT 1390  
QY 1288 TTCAATATGTTGAGCGCTTTACACCAAAAGGAGACTCAACACATTTGACGTTGGTGAGA 1347  
Db 1391 TTCAACATTTATCCAGTCCATCCACCAAGAAAGAAATGAAGATGTTTACCAGCTGGTTTGA 1450  
QY 1348 GAATCGGGATTCATGATCTTACATTTACCGCCAGCGGCTGTGGAATGATTTCTCA 1407  
Db 1451 GATTCGGGTTGCGCACTATTCACCTTTCGCTCGGGAGAGCGCGTGAATTTCTACTTTTA 1510  
QY 1408 GTGGCGGTTAGTATGTTTGAGCCAGAAATTCGCTGCTTGTAGAAATTCCTATGCCAAGACT 1467  
Db 1511 GTAGCGCGGGACCTATGAAACCCAGATGCAAAATGCAAGGTTCTCTTTTCAAAAGTG 1570

QY 1468 TCTTGCTCGCAGTTATTCTAGACGATCTTTAGCACACCCACGATCTCTGGATGATCTT 1527  
Db 1571 GCATGCTTCGACACTGTTCTTGACGATATGTATGACACTTATGGAACCCCTAGATGAATTG 1630  
QY 1528 AAATGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGCTGGATAGCTGTTCCGGAT 1587  
Db 1631 AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCTTTACAGAAAACCTTCCAGAC 1690  
QY 1588 AATCAGTTCAAAAGTTTGTCTTAGGGCTGTACAAACACAGTGAATGGAATTTGAAAAGAT 1647  
Db 1691 TAT---ATGAAACTATGTTTACCAATCTATTATGACATAGTTTCAAGAGTGGCTTGGGAG 1747  
QY 1648 GGACTCAAGGAACAGGCGGTGATGCTGGGCTATCTTCGAAAAGATGATGGGAGGGCTTG 1707  
Db 1748 GCAGAGAAGAACAGGGGCGTGAATTTGGTCAGCTTTTTCAGAAAAGGATGGGAGGATAT 1807  
QY 1708 CTCGCATCTGATACCAAGAAAGCCGAATGTTGGCAGCAAAAGTATGTGCCGACATTTCAAC 1767  
Db 1808 CTTCTGGGTTTATATGAAGAAGCTGAATGTTAGCTGCTGAGTATGTGCTCCTCTTGGAC 1867  
QY 1768 GAATATGTGGAATAATGCCAAAAGTGTCCATAGCACTTTGGCAGAGTCTGTATCTAAATC 1827  
Db 1868 GAGTACATAAAGAAATGGAATCACAATCTATCGGCCAACGTATATCTTCTGTTGAGTGAGTG 1927  
QY 1828 TTTTTCAC---TGGAGAAATTACTTCTCTGATTACATTTTACAGCAAGTAGACCTTCGGTCC 1884  
Db 1928 TTGATAAATGGAATGGGCAACTCTCTTTCGCAAGAGGCAATTAGAGAAAGTAGATTATCCAGGA 1987  
QY 1885 AAATTTCTGC-----ATCTTGTGCTTTTGTGACTGGACGACTAATCAATGACACCAAGACT 1938  
Db 1988 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCGCTGGCGGATGACAGAGACA 2047  
QY 1939 TACCAAGCCGAGAGAAACCGTGTGTAATTTGGTTTTCAGCGGTACAGTGTCTACATGAGGGAA 1998  
Db 2048 TATAAAGCTGAGAAGGCTCGTGGAGAATTTGGGCTCAGCATTTGAAATGTTTACATGAAAGAC 2107  
QY 1999 AATCCGAGTGCACAGAGGAAGAGCTCTAAGTCACTGTTTATGATATCATCGACAAAGCA 2058  
Db 2108 CATCTGTAATGTACAGAGGAAGAGGCTCTCGATCACTATATAGCAATTCGAGACCGGCG 2167  
QY 2059 CTGAAGGAATTCGAATTTGGAGTTTGGCAACCCAGCAGCAATGCCCATTTGTGTGAGA 2118  
Db 2168 GTGAGGAATCTGACAGAGAGTTTCTGAAGCCCG---ACGACGTCCCATTTGCGCTGCAG 2224  
QY 2119 AGACTGCTGTTCAACAATCGAAGAGTATGCAAGTGTGAGTCTTTTATATGATGACAGATGCTTT 2178  
Db 2225 AAGATGCTTTTCGAGGAGACAAGAGTGACGATGGTGATATTCGAAGGATGGAGATGGATTC 2284  
QY 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGCGCCGAACTCTTTTCGATCTCTGTGGC 2237  
Db 2285 GGTGTTTCCAAATTAGAAGTCAAAGATCATATCAAAGAGTGTCTCATTTGAACCGCTGCC 2343

RESULT 12  
AR240706  
LOCUS AR240706 2424 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 45 from patent US 6468772.  
ACCESSION AR240706  
VERSION AR240706.1 GI:27285855  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2424)  
AUTHORS Chappell, J., Manna, K.R., Noel, J. P. and Starke, C.M.  
TITLE Methods of making modified polypeptides  
JOURNAL Patent: US 6468772-A 45 22-OCT-2002;  
FEATURES Location/Qualifiers  
1..2424  
/organism="unknown"  
/mol\_type="genomic DNA"  
ORIGIN

Query Match		32.3%;	Score 723.4;	DB 6;	Length 2424;
Best Local Similarity		59.8%;	Pred. No. 8.7e-204;		
Matches 1327;		Conservative	0;	Mismatches 871;	Indels 21; Gaps 6;
QY	31	ATTCATATGTCATATGATACAGCTTGGTGGCAAGAAATTCGGTCAATTCAGCGCTCTGGT	90		
DB	134	ACTCCATATGCTTATGACACAGATGGGTAGCAGGGTGCCGCCATTTGATGGCTCTGCT	193		
QY	91	GCACCCCAATTTCCCAACCGCTTCAATGATTTCTGAACAATCAACTGCCAGATGGCTCG	150		
DB	194	CGCCCGCAATTTCCCAACAGTTGACTGGATTTGAANAACAGTTTAAAGATGGTTCA	253		
QY	151	TGGGGTAGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCTC	210		
DB	254	TGGGAATTCAGTCCCACTTCTGCTGTCCGACGCTTCTTGTGCCACTCTTCTTGTGTT	313		
QY	211	CTCACTCTCAAAATATGGAATAGGGGACATTTCAAGTGCAAGAAAGGGTTGAGTTGTG	270		
DB	314	CTTGTGCTCTTAAATGGAACGTTGGGGATCTGCAAGTAGCAGGGAATGGAATTCATA	373		
QY	271	AGAAACACATGGAAGAAATGAAGACGAGCTGACAAATCACAG--GCCAAGTGGATTC	327		
DB	374	AAGAGCAATCTGGAATAGTAAAGGATGAAACCGATCAAGATAGCTTGGTAAACAGACTTT	433		
QY	328	GAGTGTGTTTCTGCAATGTTAGATGAACAAAGCTTGGATTTGGATTTCTCTTAT	387		
DB	434	GAGATCATATTTCTCTCTGTTAAGAGAGCTCAATCTCTGCGCCTCGGACTTCCCTAC	493		
QY	388	CACCTCCCTTCTCATCTCCCAATCCACAAAGCGCCAGAAAGCTTCAAAAGATTCCTC	447		
DB	494	GACCTGCTTATATACATCTGTTGCAGACTAAACGGCAGGAAGATTAGCAAAACTTTCA	553		
QY	448	CTCAATGTTCTTCAACCATCAGACGCGTGTCTACTCTCTGAGGGTTTGCAGAT	507		
DB	554	AGGAGGAAATTTATGCGGTTCCGTCGCCATTTGTTGTTATCTTTAGAGGGAATCAAGAT	613		
QY	508	GTGGTGNACTGCGAGAGATCACAACTCTCAATCAAGAGCGGATCATTTTTAGCTTCC	567		
DB	614	ATAGTTGAATGGGAACGAATATGGAAGTTTAAAGTCAAGATGGTCTTTCTTAAGCTCA	673		
QY	568	CCTGCATCTACTGCTGTGTCTTCATGCACACTCAAAACAAACGATGCTCCACTTCTC	627		
DB	674	CCTGCTCTACTGCTGCGTTTTCATGCACACAGGAGACGCGAAATGCTTGAATTTCTG	733		
QY	628	AACCTGCTGCTCAGCAAAATTTGGCACTACGTTCTGCTCCATTCACCTTGTATCTT	687		
DB	734	AACAGTGTGATGATCAAGTTTGGAAATTTGTTTCCCTGCTGTATCTCTGTGATCTGCTG	793		
QY	688	GAAACCTCTGCGGCTGTGATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAA	747		
DB	794	GAAACCTCTGCTGATGATGATATTTGATGATGATGATGATGATGATGATGATGATGATG	853		
QY	748	GAAATCAAGAAATCTCTGATTTAGGTTATAGGTTATAGGTTATAGGTTATAGGTTATAG	807		
DB	854	GAAATCAAGAAAGCTTCTGATTTATGTTTACAGGCAATGGAA--CGAAAGGGAATTTGG	910		
QY	808	TGGGCAAGATGCAATCTTCTGATGTCGATGACACTGCCATGGGTCTTAGAATCTCTG	867		
DB	911	TGGGCAAGATGCAATCTTCTGATGTCGATGACACTGCCATGGGTCTTAGAATCTCTG	970		
QY	868	AGACTTCATGATATCAATGTATCTTCAAGATTTCTGAGAAATTTTCAGACGAGAAAGGA	927		
DB	971	CGGCTCATAGGTACAAATGATCTCCAGCCATTTTGAACACTTCAAGATGCCAATGGG	1030		
QY	928	GACTTCTTTTCTGCTTCCCGTCAACCGCAATTTGGTGTGACCGGATATCTTAACCTTTAT	987		
DB	1031	AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGATGTGATCTTTAT	1090		
QY	988	AGATGTTCAACAGTATGTTTTCGGGAGAAAGATTAATGGAAGAGCTAAGACTTCACT	1047		
DB	1091	AGAGCTTCCAGCTCGCANTTTCCCGGAGAAACATTTCTGATGAAGCTAAAGCTTCGCT	1150		
QY	1048	ACAAATCATCTCCAAATATGCTCTTGGCCAAACAAACGCAATTTGATAGTGGGCTGTCAAG	1107		

DB	1151	ACTAATAATTTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAAACAA	1210		
QY	1108	AGGATCTTCTCTGAGAGGTGGATGATGCTATTAAGATATCCGTGGCATAGAGTATGCCA	1167		
DB	1211	CMAAACCTGAGCCAGAGATCAAAATCGCGCTGAAGACTTCTTGGCATGCGAGTGTCCG	1270		
QY	1168	AGATTCGAGCAAGAAGTTTACATAGAGCAATTTTGGATCAAAATGATGTCTGGCTGGGAAG	1227		
DB	1271	AGAGTGAAGCAAGAGATGACTCTCAAGTGTATCGCCAGATTTATGACGATAGCAAAA	1330		
QY	1228	ACTGTGTATGAAGTGTATATGAGCAACGAAATAATTTGGAGCTGGCCAAATTCGAC	1287		
DB	1331	TGCGTTTACAAGTACCTTACCTAGCTGAACAATGAAAGATTTTATAGAGCTGGGAAAATTAGAT	1390		
QY	1288	TTCAATATGCTGAGGCTTACACCAAAAGAGACTCAACACATTTGTGAGCTGGTGAGA	1347		
DB	1391	TTCAACATTTATCAGTCCATCCCAAGAAATGAAGATGTTTACAGCTGGTTAGA	1450		
QY	1348	GAATCGGGAATTCATATGATCTTCAATTCACCCGCGAGCGCTGTGGAAATGATTTCTCA	1407		
DB	1451	GAATCGGCTTGCACATTTTACCTTCCGTCGGGAGAGCGCTGGAATTTCTACTCTTA	1510		
QY	1408	GTGCGGTGATGATGTTTGAAGCAAGATTCGCTGCTTGTAGAAATTCCTATGCCAAGACT	1467		
DB	1511	GTAGCGCGGGGACTATGAACCCCAAGTATGCCAAATGCAAGTTCCTCTTTACAAAGTG	1570		
QY	1468	TCTTGCTCGCAGTTTATCTAGAGCATTTTACGACACCCACCGATCTCTCGATGATCTT	1527		
DB	1571	GCATGCTTGCAGACTGTTCTGAGCATATGATGACACTTATGGAACCTTAGATGAATG	1630		
QY	1528	AAATGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTCTGCTGGATAGCTGGGAT	1587		
DB	1631	AAGCTATTTCACTGAGGCTGTGAGAGATGGGACCTCTCTTTACAGAAAACCTTCAAGAC	1690		
QY	1588	AATCAGTTGAAGTTTCTCTTAGGCTGTACAAACACAGTGAATGGATTTGGAAGAAT	1647		
DB	1691	TAT---ATGAACATATGTTTACCAATCTTATGACATAGTTTACAGAGTGGCTGGGAG	1747		
QY	1648	GGACTCAAGGAACAGCGCTGATGTCTGGGCTATCTTGGAAAGTATGGAGGGCTTG	1707		
DB	1748	GCAGAGAAGGAACAGGGCGTGAATTTGGTCACTTTTTCAGAAAGGGATGGAGGATAT	1807		
QY	1708	CTCGCATCGTATACCAAGAGCCCAATGCTCGGAGCAAAAGTATGTGCCGACATTTCAAC	1767		
DB	1808	CTTCTGGGTTATATGAAGAGCTGAATGCTGCTGCTGATGATGCTGCTACCTTGGAC	1867		
QY	1768	GAATATGTGAAAATGCAAAAGTGTCCATAGCACTTGGGACAGTGTACTAACTCAATC	1827		
DB	1868	GAGTACATAGAATGGAATCACATCTATCGGCCAACGTATATCTTCTGTTGAGTGGAGTG	1927		
QY	1828	TTTTTTCAC---TGGAGAAATTAATCTCTGATTAATTTTACAGCAAGTACCTTCGCTCC	1884		
DB	1928	TTGATAATGGAATGGCACTCTCTTTCGCAAGAGGCAATTAGAAGATGATTAATCAGGA	1987		
QY	1885	AAATTTCTGC-----ATCTTGTCTTTTACTGAGCACTAATCAATCAACACCAAGACT	1938		
DB	1988	AGACGTTTCTCACAGAGCTGAATAGCTCTATTTCCGCTGCGGATGACACGAAGACA	2047		
QY	1939	TACCAAGCCGAGAGAAACCGTGTGAAATTTGGTTTCCAGGCTACAGTGTCTATGAGGAA	1998		
DB	2048	TATAAGCTGAGAAGCTCTGTGAGAAATTTGGCTCCAGCATTTGAATGTTTATCAATGAAGAC	2107		
QY	1999	AATCCGAGTGCACAGAGAGAGCTCAATGATGATGATGATGATGATGATGATGATGATGATG	2058		
DB	2108	CATCTGAAATGATACAGAGAGAGGCTCTCGATCACATCTATAGCATTTCTGGAGCCGCG	2167		
QY	2059	CTGAAGGAATTAATTTGGGAGTTGGCCCAACCCAGCGAGCAATGCCCATTTGTGTGAGA	2118		
DB	2168	GTGAAGGAATGACAGAGAGATTTCTGAAAGCCG---ACGACGTCCCATTTCCGCTGCAAG	2224		
QY	2119	AGACTGCTGTTCACACTGCAAGAGTATGACAGTGTGTTTATATGATACAGAGATGGCTTT	2178		

Db 2225 AAGATGCTTTTCAGGAGACAAGAGTGACGATGGTGATATTCAAGGATGGAGATGGAATTC 2284

Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCAATGTACGCCGAACTCTTTTCGATCCTGTGGC 2237

Db 2285 GGTGTTTCCAAATTAGAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCC 2343

RESULT 13

AR266981

LOCUS AR266981 2424 bp DNA linear PAT 10-APR-2003

DEFINITION Sequence 45 from patent US 6495354.

ACCESSION AR266981

VERSION AR266981.1 GI:29696436

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2424)

AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.

TITLE Syntheses

JOURNAL Patent: US 6495354-A 45 17-DEC-2002;

FEATURES

Location/Qualifiers

source 1..2424

/organism="unknown"

/mol\_type="genomic DNA"

ORIGIN

Query Match 32.3%; Score 723.4; DB 6; Length 2424;

Best Local Similarity 59.8%; Pred. No. 8.7e-204;

Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

Qy 31 AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATCCGTCAATTGACGGCTCTGGT 90

Db 134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGNGGGTCCCGCCATTTGATGGCTCTGCT 193

Qy 91 GCACCCCAATTTCCCAACGCTTCAATGGATTCTGAACAATCAACTGCCAGATGGCTCG 150

Db 194 CGCCCGCAATTTCCCAACGCTTCAATGGATTCTGAACAATCAACTGCCAGATGGCTCG 253

Qy 151 TGGGTGAGGAGTGCAATTTTCTGCGTATGACAGAGTTTAAACACTCTCGCCCTGCTC 210

Db 254 TGGGCAATTCAGTCCCACTTTCTGCTGCTCCGACCGTCTTCTTGGCCACTCTTTCTTGTGT 313

Qy 211 CTCACCTCCTCAAAATATGGAATGAAGGCGACATTCAGTGCGAAGAAAGGGTGTGATTTGTG 270

Db 314 CTTGTGCTCTTAAATGGAACTGTTGGGATCTGCAAGTAGACAGGGAATTCGAATTCATA 373

Qy 271 AGAACAACATGGAAGAAATGAAGACGAGTCAATCAACAG--GCCAAGTGGATTC 327

Db 374 AAGCAATCTGGAACCTAGTAAGGATGAACCGATCAAGATAGCTTGGTAAACAGACTTT 433

Qy 328 GAGGTGCTGTTTCTCGCAATGTAGATGAAGCAAAAGCTTGGGATTTGGATCTTCTTTAT 387

Db 434 GAGATCATATTTCTTCTCTGTTAAGAGAGTCAATCTCTGCGCTCGGACTTCCCTAC 493

Qy 388 CACCTCCCTTCAATCTCCCAAAATCACCAGAGGCGAGAAAGCTTCAAAAGATTCCTC 447

Db 494 GACCTGCCCTTATATACATCTGTTGACAGATAAACGCGAGGAAGATTAGCAAACTTTCA 553

Qy 448 CTCATGTTCTTCATAACCATCAGACGGGTTGCTCTACTCTCTCGAGGTTTGAAGAT 507

Db 554 AGGAGGAAATTTATGCGGTTCCGTCGCCAATTTGTTATTTTATAGAGGGAATACAAGAT 613

Qy 508 GTGTGGGATGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCAATTTTAAAGCTCC 567

Db 614 ATAGTTGAATGGAACGAATATAGAGTTCAAGTCAAGATGGGTCTTCTTAAGCTCA 673

Qy 568 CCGCATCTACTGCTGTGCTTCAATGCACTCAAAACAAACAGATGCTCCTCACTTTCTC 627

Db 674 CCGCTTCTACTGCTGCTTTTCAATGCACTCAAAACAAACAGATGCTCCTCACTTTCTC 733

Qy 628 AACTTCGTGCTCAGCAAAATTTGGCGACTACGTTCTCTTGGCATTAACCACTTGATCTATT 687

Db 734 AACAGTGTGATGATCAAGTGTGGAATTTTGTTCCTGCTGATCTCTGTGGATCTGCTG 793

Qy 688 GAAACCCCTCTGGGCTGTGATACAGTGTGAACCGCTTGGGGAATCGATCGCTATTTCAGAAA 747

Db 794 GAAACCCCTCTGTGATCGTAGAATAATATGTGACGCTTGGGAATCTATAGACACTTTGAAAAG 853

Qy 748 GAAATCAAGAATCTCTGATTTACGTTTATAGTACTGGGACGCGAAGAGCGGTGGGA 807

Db 854 GAAATCAAGGAAGCTCTTGATTTATGTTTACAGGCATTTGGA--CGAAAGAGGAATTTGGG 910

Qy 808 TGGGCAAGATGCAATCTTCTGATGTGATGACACTGCGCATGGGTCTTGAAGATCTCTG 867

Db 911 TGGGCGACAGATAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTGATGCTT 970

Qy 868 AGACTTCATGAGATACAATGTATCTTCAGATGTTCTGGAGAAATTTCAAGACGAGAAAGGA 927

Db 971 CGGCTGCATAGGTACAATGTATCTCCAGCCATTTTGAACAATTTCAAGATGCAATGGG 1030

Qy 928 GACTTCTTTTGTGCTTGGCGGTCAACCGCAATTTGGTGTGACCGATATCTTAACCTTTAT 987

Db 1031 AAATTCATTTGCTCGACCGGTCAATTTCAACAAGATGTAGCAAGCATGCTGAATCTTTAT 1090

Qy 988 AGATGTTTCAAGATATGTTTTCGGGAGAAAAGATAATGGAAGAAGCTAAGACCTTCACT 1047

Db 1091 AGAGCTTCCAGCTCGCATTTCCCGAGGAACAATCTTGTATGAAGCTAAGAGCTTCGCT 1150

Qy 1048 ACAATATCATCTCAAAATGCTTTCGCCAAAACAACGCAATTTGATAAGTGGGCTGTCAAG 1107

Db 1151 ACTAAATATTTGAGAGAAAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAACAACAAA 1210

Qy 1108 AAGGATCTTCTGGAGAGTGGAGTATGCTATATAAGTATCCGTCGCATAGAGTAGTGCCA 1167

Db 1211 AAAACCTGAGCAAGAGATCAATACGCGCTGAAGACTTCTTGGCATGCCAGTGTTCG 1270

Qy 1168 AGATTGGAGGCAAGAAAGTTTACATAGACAAATTTGGATCAAAATGATGTCTGCTGGGGAAG 1227

Db 1271 AGAGTGGAGCAAGAGATCTGTCAAGTGTATCGCCAGATTTGACGCAATAGCAAAA 1330

Qy 1228 ACTGTGTATAGATGCTATATGTGAGCAACGAAAATATTTGGAGTGGCCAAATTTGGAC 1287

Db 1331 TCGCTTACAAAGCTACCTTACGTAACAATGAAGAAAGTTTATAGAGTGGGAAAATTAGAT 1390

Qy 1288 TTCATATGTTGCGAGGCTTACACCAAGAGAGCTCAACATTTCTCAGCTGTTGGAGA 1347

Db 1391 TTCATATTTACAGTCCATCCACCAAGAAAGAAATGTTTACCAAGCTGGTTTGA 1450

Qy 1348 GAATCGGATTCATATGATCTTACATTTACCCCGAGCGCTGTGGAATGTTATTTCTCA 1407

Db 1451 GATTCGGGTTGCCACTATTACCTTCGCTCGGAGAGCCGCTGGAATTTCTACTTTTA 1510

Qy 1408 GTGCGGTTAGTATGTTTGAAGCAGAAATTCGCTGTGTTGAGAAATTCCTATGCCAAGACT 1467

Db 1511 GTAGCGCGGGGACCTATGAAACCCAGTATGCAAAATGCAAGTTCCTCTTTTCAAAAGTG 1570

Qy 1468 TCTTCCCTCGCAGTTTATCTAGACCATCTTTAGCACACCAAGGATCTCTGGATGATCTT 1527

Db 1571 GCATCTTTCAGACTGTTCTGAGCATATGTATGACATTTATGGAACCCCTAGATGAATTTG 1630

Qy 1528 AAATGTTTCTTGAAGCGGTCCGAAGATGGGATATCTCTGTGCTGGATGCGTTCGGGAT 1587

Db 1631 AAGCTATTCATGAGCTGTGAGAGATGGGACTCTCTCTTTACAGAAAACCTTCAGAC 1690

Qy 1588 AATCAGTTGAAATTTGCTTCTAGGCTGTGTAACAACAGTGAATGGATTTGGAAAGAT 1647

Db 1691 TAT---ATGAAACTATGTTTACCAATCTATTATGACATAGTTTCAAGAGTGGCTTGGGAG 1747

Qy 1648 GCACTCAAGGAACAAGCGGCTGATGTGCTGGCTATCTTCGAAAAGTATGGAGGCTTG 1707

Db 1748 GCAGAGAGGAACAAGGGGCTGAATTTGGTCACTTTTTCAGAAAAGGATGGAGGATTTAT 1807

Qy 1708 CTCGCATCGTATACCAAGAGACCGCAATTTGGTGGCAGCAAGATGTTGTCGCGCATTTCAAC 1767

Db 1808 CTTCTGGGTTTATATGAAGAAGCTGAATGTTAGCTGCTGAGTATGTGCTACCTTGGAC 1867



1348 GAATCGGGATTCAATGATCTTACATTTACCCGCCAGCGCCTGTGGAATGATTTCTCA 1407  
 Db |||||  
 1451 GATTCGGGGTCCCACTATTACCTTCGTCGGAGAGCCGCTGGAATTTCTACTCTTA 1510  
 Qy |||||  
 1408 GTGGCGGTAGTATGTTTGAAGCAGAAATTCGCTGTGTGATGATGCTTATGCTCAAGACT 1467  
 Db |||||  
 1511 GTAGCGGGGGACCTATGAAACCCAGTATGCAAAATGCAGGTTCTCTTTTCAAAAGTG 1570  
 Qy |||||  
 1468 TCTTGCTCGCATTTATTTCTAGACGATCTTTACGACACCCAGGATCTCTGATGATCTT 1527  
 Db |||||  
 1571 GCATGCTTGCAGACTGTTCTGACGATATGATGACATCTTATGGAACCTTAGATGAATG 1630  
 Qy |||||  
 1528 AAATGTTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGCTGATAGCGTTCGGGAT 1587  
 Db |||||  
 1631 AAGCTATTACAGGCTGTGAGAGATGGGACCTCTCTTTTACAGAAAACCTTCAGAC 1690  
 Qy |||||  
 1588 AATCAGTTGAAAGTTTGTCTCTAGGGCTGTACAAACAGTGAATGGAATTTGAAAAGAT 1647  
 Db |||||  
 1691 TAT--ATGAAACTATGTTTACCAATCTATTATGACATAGTTTACAGAGTGGCTTGGGAG 1747  
 Qy |||||  
 1648 GGACTCAGGAACAAGGCGGTGATGTCTGGGCTATCTTCGAAAGTATGGAGGGCTTG 1707  
 Db |||||  
 1748 GCAGAGAAAGAACAGGGCGGTGAATTTGGTCAGCTTTTTCAGAAAAGGATGGAGGATAT 1807  
 Qy |||||  
 1708 CTCGCATGTTATACCAAGAGCCGAATGCTGGCAGCAAAATGATGTGCCGACATTTCAAC 1767  
 Db |||||  
 1808 CTTCTGGGTTATTTGAAGAAGCTGAATGGTTAGCTGTGATGATGTGCTACCTTGGAC 1867  
 Qy |||||  
 1768 GAATATGTGGAATGCCAAAGTGTCCATAGCATTGGCAGACAGTCTGTAACACTCAATC 1827  
 Db |||||  
 1868 GAGTACATAAGNATGGAATCACATCTATCGGCCAACGTATACTTCTGTGTGAGTGAGTG 1927  
 Qy |||||  
 1828 TTTTTCAC---TGGAGAAATTAATCTCTGATTAATTTTACAGCAAGTATAGACCTTCGGTCC 1884  
 Db |||||  
 1928 TTGATAATGGATGGGCAACTCTCTTTCGCAAGAGGCAATTAGAGAAAGTATATCCAGGA 1987  
 Qy |||||  
 1885 AAATTTCTGC-----ATCTTGTGCTTGTGACTGGACGACTTAATCAATGACACCAAGACT 1938  
 Db |||||  
 1988 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCCGCCTGGCGGATGACACGAAGACA 2047  
 Qy |||||  
 1939 TACCAGGCGGAGAGAAACCGTGGTGAATGGTTCACAGCTACAGTGTCTACATGAGGGA 1998  
 Db |||||  
 2048 TATAAGCTGAGAGGCTCGTGGAGAAATGGGCTCCAGCATTTGAATGTTACATGAAGAC 2107  
 Qy |||||  
 1999 AATCCGAGTGCACAGAGAGAAAGCTCTAAGTCAATGTTTATGATGATCATCGAACCGA 2058  
 Db |||||  
 2108 CATCTGAATGTACAGAGAGAGGCTCTCGATCATCATCTATAGCATTTCTGAGCGCGCG 2167  
 Qy |||||  
 2059 CTGAGGAATTTGAATTTGGAGTTGGCCAAACCCAGCGACATGCCCATTTGTGTGAGA 2118  
 Db |||||  
 2168 GTGAAGGAACTGACAAGAGAGTTTCTGAAGCCCG---ACGAGCTCCCATTTGCCCTGCAAG 2224  
 Qy |||||  
 2119 AGACTGCTGTTTCAACACTGCAAGAGTGTGAGCTGTTTATATGATGACAGAGATGGCTTT 2178  
 Db |||||  
 2225 AAGATGCTTTTCGAGGACACAGAGTGTGATGATGATGATGATGATGATGATGATGATG 2284  
 Qy |||||  
 2179 GGTATCTCTGCAAAAGAGATGAAAGACCATGTGACCGCAACTTTTTCGATCTGTGGC 2237  
 Db |||||  
 2285 GGTGTTTCCAAATTAAGAGTCAAGATCATATCAAAAGAGTGCTCATTTGAACCCGTGCC 2343

RESULT 15  
 AR338473  
 LOCUS AR338473 2424 bp DNA linear PAT 17-AUG-2003  
 DEFINITION Sequence 45 from patent US 6569656.  
 ACCESSION AR338473  
 VERSION AR338473.1 GI:33725250  
 KEYWORDS  
 ORGANISM Unknown.  
 Unclassified.  
 REFERENCE 1 (bases 1 to 2424)  
 AUTHORS Chappell,J., Manna,K.R., Noel,J.P. and Starks,C.M.

TITLE Syntheses  
 JOURNAL Patent: US 6569656-A 45 27-MAY-2003;  
 FEATURES Location/Qualifiers  
 source 1..2424  
 /organism="unknown"  
 /mol\_type="genomic DNA"  
 ORIGIN  
 Query Match 32.3%; Score 723.4; DB 6; Length 2424;  
 Best Local Similarity 59.8%; Pred. No. 8.7e-204;  
 Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;  
 Qy 31 AATCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCCTCAATTCACGCTCTGCT 90  
 Db 134 ACTCATCTGCTTATGACACAGATGGGTAGGAGGGTCCCGCCCATGATGCTCTGCT 193  
 Qy 91 GCACCCCAATTTCCCAAAACGCTTCAATGGATTTCTGAACAAATCAACTGCAGATGCTCG 150  
 Db 194 CGCCCGCAATTTCCCAAAACAGTTGACTGGATTTTGA AAAACAGTTAAAGATGCTTCA 253  
 Qy 151 TGGGTGAGGAGTGCATTTTCTGCGGTATGACAGAGTTTAAACACATCTCGCTCGCTC 210  
 Db 254 TGGGAAATTCAGTCCCACTTTCTGCTGTCGACCGCTCTTCTTCCCACTCTTCTTCTG 313  
 Qy 211 CTCACCTCTCAAAATATGGAATTAAGGCGACATTTCAAGTGCAGAAAGGGCTTGAGTTG 270  
 Db 314 CTTGTGCTCTTAAATGGAACGTTGGGATCTGCAAGTAGACAGGGAATGAATTCATA 373  
 Qy 271 AGAAAACATCGAAGAAATGAAGACGAAGCTGCAATCAACAG---GCCAAGTGATTC 327  
 Db 374 AAGCAATCTGGAATAGTAAGAGATGAAACCGATCAAGATAGCTTGTGTAACAGACTTT 433  
 Qy 338 GAGTGTGTTTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGATCTTCTTAT 387  
 Db 434 GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC 493  
 Qy 388 CACTCCCTTTTCTCTCCCAATCCCAAGGCGGCAAGAAAGCTTCAAAAGATTTCC 447  
 Db 494 GACCTGCTTATATACATCTGTTGCAGACTAAACGGCAGGAAGATTTAGCAAACTTTCA 553  
 Qy 448 CTCATGTTCTTCAATACATCAGACGGCTTGTCTACTCTCTGAGGCTTTGCAAGAT 507  
 Db 554 AGGAGGAAATTTATGCGGTTCCGTGCGCATTTGTTGTTATTTCTTAGAGGGAATACA 613  
 Qy 508 GTGTGGAATGGAAGAGATCAAAATCTTCAATCAAGAGACGATCATTTTAAAGCTCC 567  
 Db 614 ATAGTTGAATGGGAACGAATAATGGAAGTTCAAGAGTCAGAGTGGGTCTTCTTAA 673  
 Qy 568 CTGCACTCTACTGTTGTCTTCTCATGCACATCAAAACAAACGATGCCCTCCACTTCTC 627  
 Db 674 CCGTCTTCTACTGCTGCTGTTTTCATGCACACAGAGACGCGAATTTGAAATTTCTTG 733  
 Qy 628 AACTTCGTGCTCAGCAATTTGGCGACTACGTTCTTGGCATTAACCACTTGCATCTATT 687  
 Db 734 AACAGTGTGATCATCAAGTTTGGAAATTTTGTCTCTGCTGTATCTCTGTGATCTGCTG 793  
 Qy 688 GAAAGCTCTGCGCTGTGATACAGTTGAAAGCTTTGGGAATCGATGCTTATTTCAAGAAA 747  
 Db 794 GAAAGCTGTGATCGTAGATAATTTGTAGCCCTTGAATCTATAGACACTTTTGAAGAG 853  
 Qy 748 GAAATCAAGAAATCTCTGATTTAGGTACTGGGACGCGGAAAGAGCGGTGGGA 807  
 Db 854 GAAATCAAGGAGCTCTTGTATTTTACAGGCAATTTGAA---CGAAAGAGGAATTTGG 910  
 Qy 808 TGGGCAAGATGCAATCTTATCTCATGTGATGACACTGCGCATGGCTCTTGAATCTCTG 867  
 Db 911 TGGGCAAGCTTAATCCATAGCAGATCTTGAACACATGCTTTGGGATTTGATTTGCTT 970  
 Qy 868 AGACTTCAATGATCAATGATCTTCTGAGATGTTCTGGAATTTCTGAGACGAGAAAGGA 927  
 Db 971 CGGCTGCATAGTACAAATGATCTCCAGCAATTTTGGACAACTTCAAGATGCCAATGG 1030  
 Qy 928 GACTCTTTTGTGTTGGCGGTCAAAAGCGGAATTTGGTGTGACCGGATTAATCTTAACCTTAT 987

Db	1031	AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1090	2108	CATCCTGAATGTACAGAGGAGGCTCTCGATCACATCTATAGCAATCTTGAGCCGCG	2167
Qy	988	AGATGTTCAAGTATGTTTCCGGGAGAAAGATAATGGAAGCACTAAGACCTTCACT	1047	2059	CTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGCAATGCCCATTTGTTGAGA	2118
Db	1091	AGAGCTTCCAGCTCGCATTTCCGGGAGAAACATTTCTTGATGAAGCTAAAGCTTCGCT	1150	2168	GTGAAGGAATCTGACAAAGAGAGTTTCTGAAGCCCG---	2224
Qy	1048	ACAAATCATCTCCAAATGCTCTTGCCAAACCAACGCAATTTGATAGTGGGTGTCAG	1107	2119	AGACTGCTGTTCAACACTGCAAGAGTGAATGAGCTGTTTTATATGTACAGAGATGGCTTT	2178
Db	1151	ACTAATATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAACAAACA	1210	2225	AGATGCTTTTCGAGGAGCAAGGTGACGATGGTGATATTTCAAGGATGGAGATGGATTC	2284
Qy	1108	AGGATCTTCCCTGGAGAGTGGAGTATGCTATPAAAGTATCCGTCATAGAGTATGCCA	1167	2179	GGTATCTCTGACAAAGAGATGAAAGACCATGTACGCCGAACTCTTTTCGATCCTGTGGC	2237
Db	1211	CAAAACCTGAGCCCAAGAGATCAAAATCGCGCTGAGACTTCTTTGGCATGCCAGTGTCCG	1270	2285	GGTGTTCCAAATTAGAAGTCAAAAGATCATATCAAGAGATGTCTCTCATTTGAACCCGCTGCC	2343

Search completed: August 24, 2004, 20:44:29  
Job time : 8703.94 secs

Db	1031	AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1090	2108	CATCCTGAATGTACAGAGGAGGCTCTCGATCACATCTATAGCAATCTTGAGCCGCG	2167
Qy	988	AGATGTTCAAGTATGTTTCCGGGAGAAAGATAATGGAAGCACTAAGACCTTCACT	1047	2059	CTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGCAATGCCCATTTGTTGAGA	2118
Db	1091	AGAGCTTCCAGCTCGCATTTCCGGGAGAAACATTTCTTGATGAAGCTAAAGCTTCGCT	1150	2168	GTGAAGGAATCTGACAAAGAGAGTTTCTGAAGCCCG---	2224
Qy	1048	ACAAATCATCTCCAAATGCTCTTGCCAAACCAACGCAATTTGATAGTGGGTGTCAG	1107	2119	AGACTGCTGTTCAACACTGCAAGAGTGAATGAGCTGTTTTATATGTACAGAGATGGCTTT	2178
Db	1151	ACTAATATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAACAAACA	1210	2225	AGATGCTTTTCGAGGAGCAAGGTGACGATGGTGATATTTCAAGGATGGAGATGGATTC	2284
Qy	1108	AGGATCTTCCCTGGAGAGTGGAGTATGCTATPAAAGTATCCGTCATAGAGTATGCCA	1167	2179	GGTATCTCTGACAAAGAGATGAAAGACCATGTACGCCGAACTCTTTTCGATCCTGTGGC	2237
Db	1211	CAAAACCTGAGCCCAAGAGATCAAAATCGCGCTGAGACTTCTTTGGCATGCCAGTGTCCG	1270	2285	GGTGTTCCAAATTAGAAGTCAAAAGATCATATCAAGAGATGTCTCTCATTTGAACCCGCTGCC	2343
Qy	1168	AGATTCGAGCAAGAGTATACATAGAGCAATTTGGATCAATGATGTCTGGCTGGGGAAG	1227			
Db	1271	AGAGTGAAGCAAGAGATACTGTCAAGTGTATGCCAGATTTATGCACGCATAGCAAAA	1330			
Qy	1228	ACTGTGTATAAGATGCTATATGTAGCAACGAAATATTTGGAGCTGGCCAAATTTGGAC	1287			
Db	1331	TGGGTTTACAAGCTACCTACGTGMAACAATGAAAGTTTTTTAGAGCTGGGAAAATTAGAT	1390			
Qy	1288	TTCAATATGCTGAGGCCCTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGTGGAGA	1347			
Db	1391	TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAAGATGTTACAGCTGTGTTAGA	1450			
Qy	1348	GAATCGGGATTTCAATGATCTTACATTCACCCCGCAGCGGCTGTGGAATGTATTTCTCA	1407			
Db	1451	GATTCGGGTTGCCACTATTACCTTCGCTCGGAGAGGCGCTGGAAATCTACTTCTTA	1510			
Qy	1408	GTGCGGTTAGTATGTTTGGAGCCAGAAATTCGCTGCTGTGATGATTTGCTATGCCAAGACT	1467			
Db	1511	GTAGCGCGGGGACTATGAACCCAGTATGCCAAATGCAAGTTTCCCTTTTACAAAAGTG	1570			
Qy	1468	TCCTGCTCGCAGTATTCTAGAGACTTTTACGACACCCACCGATCTCTGGATGATCTT	1527			
Db	1571	GCATGCTTGCAGACTGTTTGGAGCAGTATGTATGACACTTATGGAACCTTAGATGAATTG	1630			
Qy	1528	AAATTTGTTCTCTGAAGCGGTCGAGATGCGATATCTGTGCTGGATAGCGTTCCGGAT	1587			
Db	1631	AAGCTATTCACTGAGGCTGTGAGAAGATGGGACCTCTCTTTACAGAAAACCTTCCAGAC	1690			
Qy	1588	AATCAGTTGAAGTTTCTCTAGGCTGTACACACAGTGAATGGATTTGGAAAAGAT	1647			
Db	1691	TAT---ATGAAACTATGTTACCAATCTATTATGACATAGTTACGAGTGGCTTGGAG	1747			
Qy	1648	GGACTCAAGGAAACAGGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGAGGGCTTG	1707			
Db	1748	GCAGAGAAGGAACAGGGCGGTGAATTTGTCAGCTTTTTCAGAAAAGGATGGAGGATAT	1807			
Qy	1708	CTCGCATCGTATACAAAGAGCCGAATGGTGGGAGCAAGATGATGCGGACATTTCAAC	1767			
Db	1808	CTTCTGGGTTATATGAAGAAGCTGAATGGTTAGCTGTGATGATGTGCTTACCTTGGAC	1867			
Qy	1768	GAATATGTGGAATGCAAGAGTGCATAGCACTTGGACATGCGACAGTCTACTAAACTCAATC	1827			
Db	1868	GAGTACATAAAGAATGGAAATCAATCTATCGCCCAACGTATACITCTGTGTGAGTGAATG	1927			
Qy	1828	TTTTTTCAC---TGGAGAATTAATCTTCTGATTAATTTTACAGCAAGTAGACCTTCGGTCC	1884			
Db	1928	TTGATAATGGATGGCACTCTTTCGCAAGAGGCAATAGAGAAAAGTAGATATCCAGGA	1987			
Qy	1885	AAATTTCTGC-----ATCTGTGTCTTTGACTGGACGACTAATCAATGACACCAAGACT	1938			
Db	1988	AGACGTGTTCTCACAGAGCTGAATAGCCCTCATTTCCCGCTGGCGGATGACACGAAGACA	2047			
Qy	1939	TACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGTCTACATGAGGAA	1998			
Db	2048	TATAAGCTGAGAGGCTGTGGAGAAATTTGGCGTCCAGCATTTGAATGTTTACATGAAGAC	2107			
Qy	1999	AATCCGAGTGCACAGAGAAAGCTCTAAGTCAATGTTTATGTTATCATCGACAACGCA	2058			

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 11:17:28 ; Search time 835.957 Seconds

(without alignments)  
11388.394 Million cell updates/sec

Title: US-10-041-007-38

Perfect score: 2241

Sequence: 1 atgtttcagtcctcaggcgca.....ttttogacctcgtggcgtag 2241

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:.\*  
1: \_geneseqn1980s:.\*  
2: \_geneseqn1990s:.\*  
3: \_geneseqn2000s:.\*  
4: \_geneseqn2001as:.\*  
5: \_geneseqn2001bs:.\*  
6: \_geneseqn2002a:.\*  
7: \_geneseqn2003as:.\*  
8: \_geneseqn2003bs:.\*  
9: \_geneseqn2003cs:.\*  
10: \_geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1078	48.1	2861	3	AAA38937
2	723.4	32.3	2424	2	AAx08654
3	723.4	32.3	2424	3	AAA38932
4	723.4	32.3	2424	4	AAf73382
5	723.4	32.3	2525	2	AAx87530
6	723.4	32.3	2528	2	AAx87529
7	723.4	32.3	2528	2	AAx87531
8	723.4	32.3	2571	2	AAx87504
9	638.2	28.5	2700	2	AAf97447
10	638.2	28.5	2700	3	AAA38931
11	450.4	20.1	2186	4	AAf73412
12	416	18.6	1890	4	AAf73421
13	412.2	18.4	2205	2	AAx08663
14	412.2	18.4	2205	4	AAf73391
15	411.2	18.3	1865	3	AAA38933
16	411.2	18.3	1885	2	AAx87532
17	411.2	18.3	1885	2	AAx87534
18	411.2	18.3	1885	2	AAx87533
19	411.2	18.3	1888	2	AAx87505
20	410.6	18.3	2196	2	AAx08643
21	410.6	18.3	2196	3	AAA38927
22	410.6	18.3	2196	4	AAf73371
23	404.4	18.0	1967	2	AAx87513

24	404.4	18.0	1967	2	AAx08655	AAx08655	Grand fir
25	404.4	18.0	1967	4	AAf73383	AAf73383	Grand fir
26	403.4	18.0	2460	3	AAZ99415	AAZ99415	Spybean c
27	403.4	18.0	2460	6	AAZ40203	AAZ40203	Soybean c
28	402.8	18.0	2089	2	AAx08645	AAx08645	Limonene
29	402.8	18.0	2089	3	AAA38938	AAA38938	Grand fir
30	402.8	18.0	2089	4	AAf73373	AAf73373	Grand fir
31	400.2	17.9	2018	2	AAx08644	AAx08644	Pinene sy
32	400.2	17.9	2018	3	AAA38922	AAA38922	Grand fir
33	400.2	17.9	2018	4	AAf73372	AAf73372	Grand fir
34	387.2	17.3	2429	3	AAf73413	AAf73413	Grand fir
35	373.6	16.7	1513	3	AAa69551	AAa69551	Pinus rad
36	373	16.6	2587	2	AAf15761	AAf15761	Gibberell
37	372	16.6	1634	3	AAa69644	AAa69644	Pinus rad
38	370.6	16.5	2013	4	AAf73411	AAf73411	Grand fir
39	347.8	15.5	1785	2	AAx87536	AAx87536	Gamma-hum
40	347.8	15.5	1785	2	AAx87535	AAx87535	Gamma-hum
41	347.8	15.5	1785	3	AAA38934	AAA38934	Grand fir
42	347.8	15.5	1977	2	AAx87506	AAx87506	Grand fir
43	346.2	15.4	1785	2	AAx87537	AAx87537	Gamma-hum
44	327.8	14.6	2784	2	AAf09940	AAf09940	Anther ea
45	306	13.7	2109	3	AAZ99417	AAZ99417	Wheat cop

## ALIGNMENTS

### RESULT 1

AAA38937

ID AAA38937 standard; DNA; 2861 BP.

XX

AC AAA38937;

XX

DT 25-AUG-2000 (first entry)

XX

DE Grand fir abietadiene synthase DNA sequence SEQ ID NO:55.

XX

XX Synthase; protein co-ordinate data; active site; modification; terpenoid; 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase; isoprenoid; breeding programme; fragrance; flavour; pheromone;

KW defensive agent; pigment; antitumour; steroid hormone;

KW signal transduction pathway; bile acid; affinity purification;

KW photoreceptor; enzymatic synthesis; nutrient supplement;

KW immunological reagent; ds.

XX

OS Abies grandis.

XX

PN WO200017327-A2.

XX

PD 30-MAR-2000.

XX

PF 17-SEP-1999; 99WO-US021419.

XX

PR 18-SEP-1998; 98US-0100993P.

PR

PR 22-APR-1999; 99US-0130628P.

PR

PR 23-AUG-1999; 99US-0150626P.

XX

XX (KENT ) UNIV KENTUCKY RES DEPT.

PA

PA (SALK ) SALK INST BIOLOGICAL STUDIES.

PI

PI Chappell J, Manna KR, Noel JP, Starks CM;

XX

XX WPI; 2000-292839/25.

DR

DR P-PSDB; AAy90858.

XX

XX Novel terpene synthase enzymes, useful for producing terpene

PT hydrocarbons, e.g. fragrances or anticancer agents, are derived from known

PT enzymes by specific amino acid alterations.

XX

XX Disclosure; Page 439-443; 450pp; English.

XX

XX The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 465-535 of a 548



CC amino acid (aa) sequence (1a), given in AAY90831. (I) contains nine alpha  
 CC -carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom ) within tabulated ranges, and have an ordered  
 CC arrangement of R groups (defining aa side chains), excluding specific  
 CC tabulated arrangements (tables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids, (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumour agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors, and as co-factor side chains.  
 CC Some syntheses with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new syntheses may produce novel terpene  
 CC products. AAA3910 to AAA38938 and AAY90831 to AAY90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 2861 BP; 890 A; 514 C; 639 G; 818 T; 0 U; 0 Other;  
 Query Match 48.1%; Score 1078; DB 3; Length 2861;  
 Best Local Similarity 68.8%; Pred. No. 0;  
 Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;  
 1 ATGTTTCAGTCCATGGGCGACGTTGAACGAATCCATCTGCATATGATACAGCTTGGGTG 60  
 378 ATGTTTATAGTATGGGCTATGGCGAAGCAATCCCTCTGCATATGACATCTTGGGTA 437  
 61 GCAGATATCCGTCATATGAGGCTCTGGTGCACCCCAATTTCCCAAAAGCTTCAATGG 120  
 438 GCAAGGATTTCCAGAGTTGATGGTCTGACACCTCACATTTCTCGAGCGTTGATGG 497  
 121 ATTCGAACATCAACTGCCAGATGGCTCGTGGGTGAGGAGTGCATTTTCTGGCGTAT 180  
 498 ATCTTCAAAATCAGTTGAAAGATGGGTCTTGGGGTGAAGGATTTCTACTTTGGCATAT 557  
 181 GACAGATTTTAAACACTCTCGCTGCTCTCACTCTCAAAATATGGAATGAAGGGGAC 240  
 558 GACAGATATCTGGCTACATTTGCATGATTAATTAACCTTACCTCTGGCTACTGGGAG 617  
 241 ATTCAAGTGCAGAAAGGGTTGAGTTGTGAGAAAAACATGGAAGAAATGAAGGACGAA 300  
 618 ACACAGTACAGAAAGGATTTGAATTTCTCAGGACACAAAGTGAAGATGAATGAA 677  
 301 GCTGACATACAGCCCAAGTGGATTCAGGTGCTGTTTCTGCAATGTTAGATGAAGCA 360  
 678 GCTGATAGTCATAGGCCAAGTGGATTTGAAATAGTATTTCTGCAATGCTTAAAGGAAGCT 737  
 361 AAAAGCTTGGATTTGGATTTCTTATCACTCTCCCTTTTCTCCCAATCCACCAAAAG 420  
 738 AAAATCTTAGGCTTGGATTTGGCTTACGATTTGGATTTCTGCAATCAATCATCGAAAG 797  
 421 GCGCAGAAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCAATAACATCAGACGGCGTTG 480  
 798 GCGGAGGCTAAGCTTAAAGGATTTCCCACTGATGTTCTTATGCTTCTTCCAAACAGTTA 857  
 481 CTCTACTCTCTGGAGGGTTGCAAGATGTTGGTGGACTGCGAGAGATCACAATCTTCAA 540  
 858 TTGTATTTCTTGGAAAGGTTTACAGAAATAGTAGACTGGCAGAAATATGAAATTTCA 917  
 541 TCAAGAGAGCGGATCAATTTTAAAGTCTCCCTGCTATCTACTGCTTGTGTTCTTCAATGCAACT 600  
 918 TCAAGGATGGATCAATTTCTAGCTCTCGGCTATCAGCGCTGTTATTCATGCGTACA 977  
 601 CAAAACAAACGATGCTCCATTTCTCAACTTCGCTCAGCAATTTGGCGCACTACGTT 660  
 978 GGGACAAAAAGTCTTGGATTTCTTGAACCTTTGCTTGAAGAAATTCGGAAACCAATG 1037  
 661 CTTTGGCATATCCACTTGCATCTATTTGAACGCTCTGGGCTGTCGATACAGTTGAACGC 720  
 1038 CTTTGTCTACTATCCGCTTGTATTTTGAACGCTTGTGTTGGCGGTGATACAGTTGAGCGG 1097

QY 721 TTGGGAATCGATCGCTATTATTTCAAGAAAGAAATCAAGAAATCTCTGATTAAGTTATAGG 780  
 DB 1098 CTAGGTATCGATCGTCAATTTCAAGAGGAGATCAAGGAAGCATTTGGATATATGTTTACAGC 1157  
 QY 781 TACTGGGACCGCGAAAGAGCGGTGGATGGGCAAGATGCATCTTATCTCTGATGTCGAT 840  
 DB 1158 CATTTGGGA---CGAAAGAGCATTTGGATGGGCGAGAGAGAAATCTTGTCTCTGATATTGAT 1214  
 QY 841 GACACTGCCATGGGTCTTAGAATCTCTAGACTTCTCATGATACAAATGATGATCTTCAGATGTT 900  
 DB 1215 GATACAGCCATGGGCTTTCGAATCTTGAGATTTACATGATACAAATGATCTCTCAGATGTT 1274  
 QY 901 CTGGAGAAATTTACAGACGAGAAAGAGACTCTTTTGTCTTTGCGGTCAAAACCAATTT 960  
 DB 1275 TTAAACAACTTTAGAGATGAGATGGGAGTCTTTTGTCTTTGGGTCAAAACACAGAGA 1334  
 QY 961 GGTGTGACCGATTAATCTTAACCTTTATAGATGTTTCAAGATGATGTTTTCGGGAGAAAG 1020  
 DB 1335 GGAGTTACAGACATGTTTAAACGTCATCTGTTTTCATGTTTTCATTTCCGGAGAAACG 1394  
 QY 1021 ATATGGAAGAGCTAAGACTTTCATCAAAATCATCTCCAAATGCTCTTTGCCAAAAAC 1080  
 DB 1395 ATATGGAAGAGCAAACTCTGTACCGAAGGTATCTGAGGAATGCTCTGAAAATGTG 1454  
 QY 1081 AACGCAATTTGATAGTGGGCTGTCAGAGGATCTTCTCTGGAGAGGTGGATATGCTATA 1140  
 DB 1455 GATGCTTTTCACAAAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGATATGCACTC 1514  
 QY 1141 AAGTATCCGTGGCATAGATGATGCCAAGATTTGAGGCAAGATTTACATAGAGCAATTT 1200  
 DB 1515 AAATATCCCTGGCATAGAGATGATCCAAAGTTTGAGGCTAGAGCTATATTTGAAAACAT 1574  
 QY 1201 GGATCAAAATGATGTCGTGGTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAA 1260  
 DB 1575 GGGCCAGATGATGTGGCTTTGGAAAACTGTATATATGATGCCATACATTTTCAATGAA 1634  
 QY 1261 AAATATTTGGAGCTGGCCAAATTTGCAATTTCAATTTGTTGAGGCTTACACCAAAAGAG 1320  
 DB 1635 AAGTATTTAGAACTAGCGAAATCTGGACTTCAATAGGTGGAGTCTATACACCAAAAGAG 1694  
 QY 1321 ACTCAACACATTTGTCAGCTGGTGGAGAAATCGGGATTTCAATGATCTTATCATTTCCCGC 1380  
 DB 1695 CTTCAAGATCTTCAAGGTGGTGGAAATCATCCGGTTTTCACGGATCTGAAATTTCACTCGT 1754  
 QY 1381 CAGCGGCTCTGGAAATGATTTCTCAGTGGGCTTAGTATGTTTGGAGCAGAAATTCGCT 1440  
 DB 1755 GAGCGTGTGACGAAATATATTTCTCACCGGCTATCTTTTATCTTTGAGCCCGAGTTTCT 1814  
 QY 1441 GCTTGTAGAAATTCCTATGCGCAAGACTTCTTGGCTCGCAGTTTATTTCTAGACGATCTTTAC 1500  
 DB 1815 AAGTGCAGAGAGGTTTATACAAAACCTTCCAAATTTTCACTGTTATTTTAGATGATCTTTAT 1874  
 QY 1501 GACACCCACGATCTCTGGATGATCTTAAATTTGTTCTCTGAAGCGGTCCGAAGATGGGAT 1560  
 DB 1875 GACGCCCATGGATCTTTAGACGATCTTAAAGTTGTTTACAGAAATCAGTCAAAAGATGGGAT 1934  
 QY 1561 ATCTCTGTGTGATAGCGTTCGGGATTAATCAAGTTGAAAGTTTCTCTAGGCTGTATC 1620  
 DB 1935 CTATCACTAGTGGAA---CCAAATGCCCAACAAATATGTTTGTGGGTTTCTAC 1991  
 QY 1621 AACACAGTGAATGGATTTTGGAAAAAGATGGACTCAAGGAAACAAGCCCGTGTATGCTGGGC 1680  
 DB 1992 AATATCTTTTATGATATAGCAAAAGAGGACGTTGAGAGCGGCGGCGATGTTGCTAGGC 2051  
 QY 1681 TATCTTTCGAAAAGTATGGGAGGGCTTGTCTCGGATCGTATACCAAGAGCCGATGCTG 1740  
 DB 2052 TACATTTCAAAATGTTTGGAAAAGTCCAACTTGAAGCTTTACACGAAAGAGCAGAAATGCTCT 2111  
 QY 1741 GCAGCAAGATATGTCGCACTTCAACGAATATGTGGAATAATGTGGAATAATGCCAAGTGTCCATAGCA 1800  
 DB 2112 GAAGCTAAATATGTGCCATCTCTTCAATGAATATCATAGAGAAATGCGAGTGTGTCAATAGCA 2171



QY 1801 CTTCGACAGTCTGTAACCTCAATCTTTTTCACCTGGAGATTACTCTCTGATTACATT 1860  
 Db 2172 TTGGGACAGTCTGTTCTCTATTAGTGTCTTTTTCACCTGGGAGGTTCTTACAGATGAAGTA 2231  
 QY 1861 TTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCATCTTTGTGTCTTTTGACTGGACGACTA 1920  
 Db 2232 CTCTCCAAATTTGATCGCGAATCTAGATTTCTTCAACTCATGGCTTTACAGGGCGTTTG 2291  
 QY 1921 ATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
 Db 2292 GTGAATGACACCAAACTTTATCAGGCAGAGAGAGGTCAAGGTGAGGTGCTTGTGCCATA 2351  
 QY 1981 CAGTGTACATCAGGGAATTCGGAGTGCACAGAGGAGAGCTCTAAGTCATGTTTAT 2040  
 Db 2352 CAATGTTATGAGGACCATCTTAAATCTCTGGAAGAAGAGCTCTACAACATGTCAT 2411  
 QY 2041 GGTATCATCGACAACCACTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGCAAT 2100  
 Db 2412 AGTGTATGGAATATGCCCTCGAAGATTTGAATAGGAGTTGTGNA-----TAACAA 2465  
 QY 2101 GCCCAATGTTGTGGAAGAGCTGCTTTCAACACTGCAAGAGTGATGAGCTGTTTAT 2160  
 Db 2466 ATACCGGATTTTACAAAGAGCTGGTTTGTAACTGCAAGATAATGCAACTCTTTTAT 2525  
 QY 2161 ATGTACAGATGGCTTTGGTATCTC---TGACAAGAGATGAAGACCATGTGACCGCA 2217  
 Db 2526 ATGCAAGGGGATGGTTTGACATCATATATGGAATTTAAGAGCATGTCAAAAT 2585  
 QY 2218 ACTCTTTTCGATCCTGTGGCGTAG 2241  
 Db 2586 TGCTCTTCCACAGTTGCTTAG 2609

RESULT 2  
 ID AAX08654 standard; cDNA; 2424 BP.  
 XX AAX08654;  
 AC AAX08654;  
 DT 27-SEP-1999 (first entry)  
 XX Grand Fir monoterpene synthase clone AG1.28.  
 XX Myrcene synthase; limonene synthase; pinene synthase; flavour;  
 KW monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;  
 KW probe; ss.  
 XX Abies grandis.  
 XX Key Location/Qualifiers  
 FH CDS 2..2350  
 FT /\*tag= a  
 FT /product= "Potential diterpene cyclase"  
 XX WO9902030-A1.  
 XX 21-JAN-1999.  
 XX 10-JUL-1998; 98WO-US014528.  
 XX 11-JUL-1997; 97US-0052249P.  
 XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
 XX Bohlmann J, Steele CL, Croteau RB;  
 XX WPI; 1999-120396/10.  
 DR P-PSDB; AAW85703.  
 XX New isolated gymnosperm monoterpene synthase DNA - obtained from Grand  
 fir (Abies grandis), used to provide plants with modified production of  
 monoterpenes, e.g. myrcene, limonene or pinene.

PS Example 1; Page 86-90; 121pp; English.  
 XX Nucleotide sequences encoding myrcene synthase, limonene synthase and  
 CC pinene synthase from Grand fir may be incorporated into any organism  
 CC (e.g. intact plant, animal, microbe), or derived cell culture that  
 CC produces geranyl diphosphate for the production of the aforementioned  
 CC enzymes or their products. The sequences when expressed in transfected  
 CC cells may also be used for the production or modification of flavour and  
 CC aroma properties, improvement of defense capability, and the alteration  
 CC of other ecological interactions mediated by myrcene, limonene, pinene,  
 CC or their derivatives. In particular they can be used for the production  
 CC of plant seeds for the extraction of oil or meal. cDNAs encoding myrcene  
 CC synthase, pinene synthase and limonene synthase from Grand Fir were  
 CC isolated and sequenced by the following procedure. Based on comparison of  
 CC sequences of different plant synthase enzymes, four conserved sequences  
 CC were identified for construction of four degenerate primers (AAX08646-  
 CC 49). One of these primers (AAX08648) was an antisense primer and used for  
 CC PCR with the other three sense primers. Only a combination of primers in  
 CC AAX08648 and AAX08649 generated a PCR product which was 110 base pairs in  
 CC length. The amplified product was used in a plasmid construct to  
 CC transform E.coli XL1-Blue cells and the inserts were sequenced. Four  
 CC probes (AAX08650-53) were devised from the sequenced insert and were used  
 CC to screen a Grand Fir cDNA library. One of these clones is the myrcene  
 CC synthase sequence given in AAX08643. The other clones identified are  
 CC described in AAX08654-56  
 XX  
 SQ Sequence 2424 BP; 730 A; 472 C; 575 G; 647 T; 0 U; 0 Other;  
 Query Match 32.3%; Score 723.4; DB 2; Length 2424;  
 Best Local Similarity 59.8%; Pred. No. 4.1e-214;  
 Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;  
 QY 31 AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATCCCGTCAATTTGACGCTCTGGT 90  
 Db 134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGGGTCCCGCCATTGATGGCTCTGCT 193  
 QY 91 GCACCCCAATTTCCCAACGGCTTCAATGGATTTGAAACAATCAATCAATGCCAGATGGCTCG 150  
 Db 194 CGCCCGCAATTTCCCAACAGCTTGACTGGATTTTGAATAAACCCAGTTAAAGATGGTTCA 253  
 QY 151 TGGGTGAGGAGTGATTTTCTGGGTATGACAGAGTTTAAACAATCTCGCTCGCTC 210  
 Db 254 TGGGAATTCAGTCCCATCTTCTGTGTCGACCGCTTCTTGGCCACTCTTCTTGTGTT 313  
 QY 211 CTCACCTCTCAAAATATGGAATAAGGCGCAGCATTTCAAGTGCAGAAAGGGTGTGAGTTTGTG 270  
 Db 314 CTGTGCTCTTAAATGGAACGTTGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA 373  
 QY 271 AGAAACACATGGAAGAAATGAAGAGCAAGCTGCAATCAAG---GCCAAGTGATTTC 327  
 Db 374 AAGAGCAATCTGGAACCTAGTAAAGGATGAACCGATCAAGATAGCTTGGTAAACAGACTTT 433  
 QY 328 GAGGTGCTGTTTCTGCAATGTTAGTGAAGCAAAAGCTTGGGATTTGATCTTCTCTAT 387  
 Db 434 GAGATCATATTTCTTCTCTGTTAAGAGAAGCTCAATCTCTGCGCTTCGACCTTCCCTAC 493  
 QY 388 CACCTCCCTTTCTCATCTCCCAATCCACAAAGCGCCAGAAAAGCTTCAAAAGATTCCC 447  
 Db 494 GACCTGCTTATATACATCTGTTGACAGCTAAACGGCAGGAAGATAGCAAACTTCA 553  
 QY 448 CTCGAATGTTTCATAAACCATCAGACGGGTTGCTCTACTCTCTGGAGGGTTGCAAGAT 507  
 Db 554 AGGGAGGAAATTTATGCGGTTCCGTCCCATTTGTTGTTATTTTAGAGGGAATACAGAT 613  
 QY 508 GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGAGGATCATTTTAAAGCTCC 567  
 Db 614 ATAGTTGAATGGGAACGAATAATGAAGTTCAAGAGTCAGGATGGGCTTTCTTAAAGCTCA 673  
 QY 568 CTGTCATCTACTGCTTGTGTTCTTTCATGACACTCAAAACAAACAGGATGCTCCACTTCTC 627  
 Db 674 CTGTTTCTACTGCTGCTGTTTTCATGCACAGGAGCGCAATGCTTGAATCTTTC 733  
 QY 628 AACTTCGTGCTCAGCAAAATTTGGCGACTACAGTTCCTTGGCATTAACCCACTTGATATT 687



PT Novel terpene synthase enzymes, useful for producing terpene  
 PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
 PT enzymes by specific amino acid alterations.

PS Disclosure; Page 414-418; 450pp; English.

XX  
 CC The present invention describes an isolated terpene synthase (I)  
 CC comprising a region with at least 20% identity to region 265-535 of a 548  
 CC amino acid (aa) sequence (Ia), given in AA90831. (I) contains nine alpha  
 CC -carbon atoms (alphaC) that have interatomic distances, between each  
 CC other, within tabulated ranges, have a centre point (within a sphere of  
 CC radius 2.3 Angstrom ) within tabulated ranges, and have an ordered  
 CC arrangement of R groups (defining aa side chains), excluding specific  
 CC tabulated arrangements (cables given in the specification). (I), and  
 CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
 CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
 CC flavours, pheromones, defensive agents, pigments, antitumor agents,  
 CC components of signal transduction pathways, precursors of steroid  
 CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
 CC Some synthases with little or no catalytic activity (and nucleic acids  
 CC encoding them) are used as controls in the analysis of products formed by  
 CC enzymatic synthesis; as nutrient supplements; for affinity purification  
 CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
 CC monitoring expression of terpene synthase or inheritance of the gene in  
 CC plant breeding programs. The new synthases may produce novel terpene  
 CC products. AA90831 to AA90838 and AA90831 to AA90859 represent  
 CC sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 2424 BP; 730 A; 472 C; 575 G; 647 T; 0 U; 0 Other;

Query Match 32.3%; Score 723.4; DB 3; Length 2424;  
 Best Local Similarity 59.8%; Pred. No. 4.1e-214;  
 Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

QY	31	AATCCATCTGATATGATACAGCTTGGTGGCAAGAAATCCGTCATTTGACGCTCTGGT	90
DB	134	ACTCCATCTGCTATGACACAGCATGGTAGCGGGTCCCGCCATTTGATGCTCTGCT	193
QY	91	GCAACCCCAATTTCCCAACGCTTCAATGATTTCTGCAACATCACTGCCAGATGGCTCG	150
DB	194	CGCCCGCAATTTCCCAACAGTTGACTGGATTTGAAAAACCACTTAAAGATGGTTCA	253
QY	151	TGGGGTAGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCCTGCCTC	210
DB	254	TGGGGGAATTCAGTCCCACTTTCTGCTGTCGACCGCTCTCTGCGCACTCTTCTTGTT	313
QY	211	CTCACTCTCAAAATATGGAATAAGGGCGACATTTCAAGTGCAGAAAGGGTTGAGTTTGTG	270
DB	314	CTTGTGCTCTTAAATGGAACGTTGGGATCTGCAAGTAGACAGGGAATGAAATTCATA	373
QY	271	AGAAAAACACATGGAAGAAATGAAGGACGAAGCTGACAAATCACAG---GCCAAGTGGATTC	327
DB	374	AAGAGCAATCTGGAACCTAGTAAAGGATGAACCGATCAAGATAGCTTTGGTAAACAGACTTT	433
QY	328	GAGTCTGTTTCTGCAATGTTAGATGAAGCAAAAGCTTTGGATTTGGATCTTCTTAT	387
DB	434	GAGATCATATTTCTCTGTTTAAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC	493
QY	388	CACCTCCCTTTCATCTCCCAATCCCAAAAGCGCCAGAAAAAGCTTCAAAAGATTTCC	447
DB	494	GACCTGCTTATATACATCTGTTGCACTTAAACGGCAGGAAGATTAGCAAACTTTCA	553
QY	448	CTCAATGTTTCTTCAATACATGACAGCGGCTTGTCTTACTCTCTCGAGGGTTTGCAGAT	507
DB	554	AGGGAGGAATTTATGCGGTTCCGTCGCCATTTGTTATTTCTTTAGAGGGAATACAGAT	613
QY	508	GTGGTGAAGTGGCAAGAGATCAAAATCTTCAATCAAGAGACGATCAATTTTAAAGTCC	567
DB	614	ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAAGTCAGGATGGGTCTTTCTTAAAGCTCA	673
QY	568	CCGTGCTACTCTGTTGTTCTCATGCACTCAAAACAAACGATGCGCTCCACTTCTC	627
DB	674	CTTGCTTCTACTGCTGCGTTTTTTCATGCAACAGGAGACGGAATGCGCTTGAATTTCTG	733

QY	628	AACCTCGTCTGACGAAATTTGGCGACTAGCTTCTTCTGCTTACCCTTACCCACTTGATCTATTT	687
DB	734	AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGTATCTCTGATCTGCTG	793
QY	688	GAACGCTCTGGGCTGTGATACAGTTGAACGCTTGGGAATTCGATCCTATTTCAAGAAA	747
DB	794	GAACGCTCTGTTGATCGTAGATAATATTTGACGCTTGGAAATCTATAGACACTTTGAAAG	853
QY	748	GAATCAAGAAATCTCTGGATTACGTTATAGGTACTGGGACGCGGAAAGAGCGGTGGGA	807
DB	854	GAATCAAGGAAGCTCTTGATTTATGTTTACAGGCAATTTGAA---CGAAAGAGGAATTTGG	910
QY	808	TGGGCAAGATGCAATCTTATCTGATGTCGATGACACTGCCATGGCTCTTAGAATCCTG	867
DB	911	TGGGGCAGACTAAATCCCATAGCAGATCTTTGAGACCACTGCTTTGGGATTTTCGATGCT	970
QY	868	AGATCTCATGATACAAATGATCTTCCAGATGTTCTTGAGAAATTTTCAGAGACGAGAAAGGA	927
DB	971	CGGCTGATAGGTACAAATGATCTTCCAGCCATTTTTCAGCACTTTCAAGATGCCANTGG	1030
QY	928	GACTTCTTTTGGCTTTCGCGTCAACGCAAAATTTGGTGTGACCGGATAATCTTAACTTTAT	987
DB	1031	AAATTCATTTGCTCGACCGCTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1090
QY	988	AGATGTTCAAGATATGTTTTCGGGAGAAAGATAATGGAAGACTAGAACCTTCACT	1047
DB	1091	AGAGCTTCCAGCTCGCATTTCCCGGAGAAACATTTCTTGATGAAGCTAAAGCTCGCT	1150
QY	1048	ACAAATCATCTCAAAATGCTTTGCCAAAAACAAGCATTTTGATAGTGGCTGTCAAG	1107
DB	1151	ACTAAATTTTGAAGAGAGCTCTTTGAGAAAGTGAGACTTCCAGTGCATGGAACAACAAA	1210
QY	1108	AAGGATCTTCTCGGAGAGGTGGATATGCTTATAAGTATCCGTGGCATAGAAATGATGCCA	1167
DB	1211	CAAACTGAGCCAGAGATCAATACGCGCTGAAGACTTCTTGGCATGCCAGTGTCCG	1270
QY	1168	AGATGGAGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGAG	1227
DB	1271	AGAGTGGAGCAAGAGATACTGTCAAGTGTATGCGCCAGATTTATGCAACGATAGCAAAA	1330
QY	1228	ACTGTGTATAGATGCTTATATGTAGCAACGAAAAATTTTGGAGCTGGCCAAATTTGAC	1287
DB	1331	TGGGTTTACAGCTACCTACGTGAACATATGAAGATTTTATAGAGCTGGGAAATTTAGT	1390
QY	1288	TTCAATATGTTGAGGCTTTTACACCAAAAGGAGACTCAACACATTTGTGAGCTGGGAGA	1347
DB	1391	TTCAACATTTACAGTCCATCCCAAGAGAAATGAAGATGTTACCAGCTGTTTATGA	1450
QY	1348	GAATCGGATTTCAATGATCTTACATTCACCCCGCAGCGGCTGTGGAAATGATTTCTCA	1407
DB	1451	GATTCGGGTTGCCACTATTTTCACTTCCGCTCGGAGAGGCGCTGGAAATTTCTACTCTTA	1510
QY	1408	GTGCGGTTAGTATGTTTGGAGCCAGATTTGCTCTGTTAGATTTGCTATGCCAAGACT	1467
DB	1511	GTAGCGGCGGAGCTATGAAACCCAGTATGCCAAATGCGAGTTTCTCTTTTACAAAAGTG	1570
QY	1468	TCTTGGCTCGCAGTATTTCTAGACGATCTTTTACGACACCCACGAGTCTCTGGATGATCTT	1527
DB	1571	GCATGCTTGACAGTGTCTTGAGACGATATGATGACACTTTATGGAACCTTAGATGAATG	1630
QY	1528	AAATGTTTCTGAGCGGCTCGAAGATGGGATATCTCTGCTGCTGGATAGGCTTGGGAT	1587
DB	1631	AAGCTATTACTAGGCTGTGAGAAAGATGGGACCTCTCTCTTTACAGAAAACCTTCCAGAC	1690
QY	1588	AATCAGTTGAAAGTTTCTTCTAGGCTGTACAAACAGTGAATGATTTTGGAAAGAT	1647
DB	1691	TAT---ATGAACTATGTTTACAAATCTATATGACATAGTTTACAGAGTGGCTTGGAG	1747
QY	1648	GGAATCAAGGAAACAGGCGCTGATGCTGGGCTTATCTTCCGAAAGATATGGGAGGCTTG	1707
DB	1748	GCAGAGAGGAAACAGGCGCTGATTTGCTGCTGCTTCTGCTGCTTCTTTCAGAAAGGAT	1807

QY 1708 CTCGCATCGTATACCAAGAGCCGAATGTCGCGACGAAGATATGCGACATTCAC 1767  
 DB 1808 CTTCTGGGTTATATGAGAGAGCTGAATGGTGTAGTGTGCTGAGTATGCGCTTACCTTGGAC 1867  
 QY 1768 GAATATGTGGAAATGCCAAGTGTCCATAGCACTTGGACAGCTGCTACTAACTCAATC 1827  
 DB 1868 GAGTACATAAGATGGATCAATCTATCGGCCAAGCTATACCTTCTGTTGAGTGGAGTG 1927  
 QY 1828 TTTTTCAC---TGAGAAATTAATCTCTCGATTAACATTTTACAGCAAGTAGACCTTCGGTCC 1884  
 DB 1928 TTGATAATGGATGGCACTCTTTTCGCAAGAGCAATTAGAGAAAGTAGATTATCCAGGA 1987  
 QY 1885 AAATTTCTGC-----ATCTTGTGCTTTGACTGACGACCTAATCAATGACACCAAGACT 1938  
 DB 1988 AGACGTGTTCTCAGAGCTGAATAGCCTCAATTTCCCGCCTGGCGGATGACACGAGACA 2047  
 QY 1939 TACCAGGCGGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTACAGTGTACATGAGGGAA 1998  
 DB 2048 TATAAGCTGAGAGGCTCGTGGAGAAATGGCGTCCAGCAATTTGAATGTTACATGAAGAC 2107  
 QY 1999 AATCGGAGTGCACAGAGAAAGCTCTAAGTCTAATGTTATGTTATGTTATCATCGACACGCA 2058  
 DB 2108 CATCTGAATGTACAGAGAAAGAGCTCTCGATCAATCTATAGCAATTTCTGAGCGCGG 2167  
 QY 2059 CTGAAGGAATGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2118  
 DB 2168 GTGAAGGAATCTGACAGAGATTTCTGAAGCCCG---ACGAGTCCCATTCGCTGCAAG 2224  
 QY 2119 AGACTGCTGTTCAACACTGCAAGAGTGTATGACGCTGTTTATATATGTAACAGAGATGGCTTT 2178  
 DB 2225 AAGATGCTTTTCAGGAGACAAAGAGTGTGACGATGGTGATATTCAGAGATGGAGATGGATT 2284  
 QY 2179 GGTATCTGTACAAAGAGATGAAGACCATGTGACCGCACTCTTTTCGATCTCTGTGGC 2237  
 DB 2285 GGTGTTTCCAAATTAGAAGTCAAGATCATATCAAGAGATGTCTCAITGAACCGTGGC 2343

RESULT 4

ID AAF73382 standard; cDNA; 2424 BP.  
 AC AAF73382;  
 XX 30-APR-2001 (first entry)  
 DT Grand fir abietadiene synthase coding sequence fragment SEQ ID NO: 15.

DE Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
 KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
 KW terpinolene synthase; insect resistance; nutrition; ss.

XX Abies grandis.

XX WO200107565-A2.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-US020264.

XX 26-JUL-1999; 99US-00360545.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;

XX WPI; 2001-182782/18.

XX P-PSDB; AAB69373.

XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
 PT synthesis in plants, e.g. for increasing resistance to pests or for  
 PT treatment of cancer.

XX Example 1; Page 119-122; 175pp; English.

XX The present invention provides the protein and coding sequences of  
 CC monoterpene synthases from the grand fir. These include (-)-camphene  
 CC synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
 CC limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
 CC and pinene synthase. The sequences can be used to produce transgenic  
 CC plants expressing high levels of the enzymes, resulting in levels which  
 CC are useful in protecting against and treating cancers, and to confer  
 CC insect resistance on plants  
 XX  
 SQ Sequence 2424 BP; 730 A; 472 C; 575 G; 647 T; 0 U; 0 Other;

Query Match 32.3%; Score 723.4; DB 4; Length 2424;  
 Best Local Similarity 59.8%; Pred. No. 4.1e-214;  
 Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

QY 31 AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAATTCGTCATATGACGGCTCTGCT 90  
 DB 134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGGGTCCCGCCCATTTGATGGCTCTGCT 193  
 QY 91 GCACCCCAATTTCCCAAAACGCTTCAATGGATTCTGAACAATCAACTGCCAGATGCTCG 150  
 DB 194 CGCCCGCAATTTCCCAAAACGCTTGAATTTTGAATAACAGTTAAAGATGTTTCA 253  
 QY 151 TGGGTGAGAGTGCATTTTCTGCGGTATGACAGAGTTTAAACACTCTCGCTGCTCCTC 210  
 DB 254 TGGGGAATTCAGTCCCACTTTCTGCTGCTCCGACCGCTCTTTTGGCCACTCTTTCTGT 313  
 QY 211 CTCACCTCTCAAAATATGAATAAGGCGGACATTTCAAGTGCAGAAAGGGGTGAGTTGTG 270  
 DB 314 CTTGTGCTCTTAAATGGAACTGTTGGGATCTGCAAGTAGACGAGGAATGAATTCATA 373  
 QY 271 AGAAACACATGGAAGAAATGAAGACGAGTGTGACAAATCAAG---GCCAAGTGAATTC 327  
 DB 374 AAGACCAATCTGAACTAGTAAAGGATGAACCGATCAAGATAGCTTGGTAAACAGACTTT 433  
 QY 328 GAGTGTGTTTCTGCAATGTTAGTGAAGCAAAAGCTTGGGATTTGATCTTCTCTAT 387  
 DB 434 GAGATCATATTTCTTCTCTGTTTAAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC 493  
 QY 388 CACCTCCCTTTCATCTCCCAAAATCCACAAAGCGCCAGAAAGAGTTCACAAAGATTCCT 447  
 DB 494 GACCTGCTTATATACATCTGTTGACAGCTAAACGCGCAGGAAAGATTAGCAAACTTCA 553  
 QY 448 CTCAATGTTCTTCATAACCATCAGACGGGTGCTCTACTCTCTGAGGGTTCGCAAGAT 507  
 DB 554 AGGAGGAAATTTATGCGGTTCGCTGCCAATTTGTTGTTATTTCTTTAGAGGAAATACAGAT 613  
 QY 508 GTGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTAAAGCTCC 567  
 DB 614 ATAGTTGAATGGGAACGAATATGGAAGTCAAGTCAGGATGGGTCTTTCTTAAAGCTCA 673  
 QY 568 CCTGCATCTACTGCTTGTCTTCATGCACTCAAAAACAAAGATGCTCCACTTCTCTC 627  
 DB 674 CTTGCTTCTACTGCTGCTGCTTTTCAATGACACAGGAGACGCGAAATGCTTGAATCTTG 733  
 QY 628 AACTTCGCTCAGCAAAATTTGCGGACTAGTCTCTTCCCTTCCCATACCCACTTGTATTT 687  
 DB 734 AACAGTGTGATGATCAAGTTTGGAAATTTGTTTCCCTGCTGATCTCTGATCTGCTG 793  
 QY 688 GAACGCTCTGGGCTGTGATACAGTTGAACGCTTGGGAATCGATCGATCTATTTCAAGAAA 747  
 DB 794 GAACGCTGTGATCGTAGATAATTTGACGCTTGGAAATCTATACACTTTGAAAG 853  
 QY 748 GAAATCAAGAAATCTCTGGATTAAGTTATAGGTACTGGGACGCGGAGAGGCTGGGA 807  
 DB 854 GAAATCAAGGAAGCTCTTGATTTATGTTTACAGGCAATTTGGAA---CGAAAGAGGAATTTGG 910  
 QY 808 TGGGCAAGTGCATCTTATCTGATGTCGATGACACTGCCATGGGCTTTAGAAATCTGT 867  
 DB 911 TGGGGCAGACTAAATCCCATAGCAGATCTTTGAGACCACTGCTTTGGGATTTGATGCTT 970  
 QY 868 AGACTTTCATGGATACAAATGTATCTTCAGATGTTTCTGAGAGAAATTTTCAGACGAGAAAGGA 927

Db 971 CGCTGATAGTACATATGTTCTCAGCCATTTTGGACAACTTCAAGATGCAATGGG 1030  
QY 928 GACTCTTTTGTTCGCGGTCAAAACGAAATTTGGTGTGACCGATTAATCTTAACCTTTAT 987  
Db 1031 AATTCATTTGCTCGACGCGTCAATTCACAAAGATGTAGCAGCATGCTGAATCTTTAT 1090  
QY 988 AGATGTTCAAGATATGTTTTCGGGAGAAAGATATGGAAGAGCTAAGACCTTCACT 1047  
Db 1091 AGAGCTTCCAGCTCGCATTTCCGGAGAAACATTTCTTGATGAAGCTAAAGCTTCGCT 1150  
QY 1048 ACAATCATCTCCAAATGCTCTTGCCAAACCAACGCAATTTGATAAGTGGGCTGTCAAG 1107  
Db 1151 ACTAATATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGCATGGAACAA 1210  
QY 1108 AAGGATCTTCTCGAGAGGTGAGTATGCTATTAAGATATCCGTGGCATAGAAGTATGCCA 1167  
Db 1211 CAAAACCTGAGCCCAAGAGATCAAAATCGCGCTCAAGACTTCTTTGGCATCGCAGTGTCCG 1270  
QY 1168 AGATTGGAGGCAAGATTTACATAGAGCAATTTGGATCAAAATGCTGTGCTGGGGAG 1227  
Db 1271 AGAGTGGAGGCAAGAGATATCTGTCAAGTGTATCGCCAGATTTGACGCGCATAGCAAA 1330  
QY 1228 ACTGTGTATAAGATGCTATATGTAGCAACGAAATATTTGAGCTGGCCAAATTTGGAC 1287  
Db 1331 TGGCTTTACAGCTACCTACGTGGAACATGAAGATTTTGTAGAGCTGGGAAATTTAGAT 1390  
QY 1288 TTCAATATGCTGAGCGCTTTACACCAAAAGGAGACTCAACACATTTGTACGCTGGTGAGA 1347  
Db 1391 TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAAGATTTTACCAGCTGTTTGA 1450  
QY 1348 GAATCGGAGTCAATGATTTCAATTCATTCACCCGCGAGCGCTGTGGAATGTATTTCTCA 1407  
Db 1451 GATTCGGGGTTCGCACTATTTTCACTTCGCTCGGAGAGGCGCTGGAATTTCTACTTTTA 1510  
QY 1408 GTGGCGTTAGTATGTTTGAAGCAGAAATTCGCTGCTGTGATGATTCGCTATGCAAGACT 1467  
Db 1511 GTAGCGCGGGACCTATGACCCAGTATGCCAATGCAAGTTCCTCTTTTACAAAGTG 1570  
QY 1468 TCTTGCTCGCAGTATTTCTAGACGATCTTTACGACACCCACCGATCTCTGATGATCTT 1527  
Db 1571 GCATGCTTCGAGACTGTTCTGAGCAGATATGATGACACTTATGGAACCTTAGATGATG 1630  
QY 1528 AATTTGCTCTGAAGGGTCCGAAGATGGGATATCTGTGCTGGATAGCGTTCGGGAT 1587  
Db 1631 AAGCTATTTCACTGAGGCTGTGAGAAGATGGGACCTCTCTTTACAGAAACCTTCCAGAC 1690  
QY 1588 AATCAGTTGAAAGTTGCTTCTTAGGCTGTACAACACAGTGAATGATTTGGAAGAT 1647  
Db 1691 TAT---ATGAACTATGTTACCAAAATCTATTATGACATAGTTCCAGGCTGCTGGAG 1747  
QY 1648 GGAATCAAGCAAGCGCTGATGCTGGGCTATCTTCAAAAGTATGGGAGGGCTTG 1707  
Db 1748 GCAGAGAGGCAAGCGGCTGATTTGGTCACTTTTTCAGAAAGGATGGGAGATTAT 1807  
QY 1708 CTGCGCATGATACCAAGAGCCGAATGCTGGCAGCAAGATATGCGCGCATTCAC 1767  
Db 1808 CTCTGGGTTATATGAAGAGCTGAATGGTGTAGTGTGCTGATGATGCTGCTTGGAC 1867  
QY 1768 GAATATGTGGAATGCAAGTGTCTCATAGCACTTGGCAGACTGCTACTAACTCAATC 1827  
Db 1868 GAGTACATAAAGAAATGGAATTCATCTATCGGCCAAGTATATCTTCTGTTGAGTGGAGT 1927  
QY 1828 TTTTTCAC---TGGAGAAATTTCTCTGATTTACATTTTACGCAAGTAGACCTTCGGTCC 1884  
Db 1928 TTGATATGATGGGCACTCTTTTTCAGAGGCAATGAGAAAGTAGATATTCAGGA 1987  
QY 1885 AAATTTCTGC-----ATCTTGTGTTTTCGACGCAATTAATATGACACCAAGACT 1938  
Db 1988 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCGCGCTGGCGATGACACGAACA 2047  
QY 1939 TACCAGGCGGAGAAACCGTGTGATTTGGTTTCCAGGCTACAGTGTCTACATGAGGAA 1998

Db 2048 TATAAGCTCAGAGGCTCGTGAGAAATTTGGCTCCAGCATTGAATGTTACATGAAGAC 2107  
QY 1999 AATCCGAGTGCAAGAGGAAAGCTCTAAGTCATGTTTATGATATCATCGACAAACGCA 2058  
Db 2108 CATCTGAAATGTACAGAGGAAAGGCTCTCGATCACAATCTATAGCATTTCTGGAGCCGCG 2167  
QY 2059 CTGAAGGAATGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTTGTGAGA 2118  
Db 2168 GTGAAGGAATGCAAGAGAGTTTTCGAAGCCCG---ACGACGTCCTCAATTCGCTGCAAG 2224  
QY 2119 AGACTGCTGTTTCAACACTGCAAGAGTGAATGAGCTGTTTATATGATACAGAGATGGCTTT 2178  
Db 2225 AAGATGCTTTTCGAGGAGCAAGTGACGATGGTATATTCAGGATGAGATGGATTC 2284  
QY 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGACGCGAACTCTTTTCGATCTCTGTGGC 2237  
Db 2285 GGTGTTTCCAAATTTAGAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCC 2343

RESULT 5  
AA87530  
ID AA87530 standard; DNA; 2525 BP.

XX AC AA87530;  
XX AC  
XX 08-OCT-1999 (first entry)  
XX DE E-alpha-bisabolene synthase computer-generated nucleic acid sequence.  
XX DE E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;  
XX KW gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour; ss.  
XX OS Synthetic.  
XX OS Abies grandis.  
XX PH Location/Qualifiers  
XX FT 1. 2454  
XX FT /\*tag= a

XX MO9937139-Al.  
XX PN 29-JUL-1999.  
XX PD  
XX PF 21-JAN-1999; 99MO-US001300.  
XX PR 22-JAN-1998; 98US-0072204P.  
XX PS (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX PI Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;  
XX WPI; 1999-458569/38.  
XX P-PSDB; AAY06567.  
XX Grand fir (Abies grandis) sesquiterpene synthase genes and protein products, useful for sesquiterpenoid production.  
XX Example 11; Page 95-98; 136pp; English.

This is the sequence of a computer-generated nucleic acid coding for a grand fir (Abies grandis) E-alpha-bisabolene synthase polypeptide (see AAY06567) that has conservative amino acid substitutions relative to the native E-alpha-bisabolene synthase sequence given in AAY06562. E-alpha-bisabolene synthase is a wound-inducible enzyme capable of generating multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl diphosphate. The invention provides gymnosperm sesquiterpene synthase nucleic acids and recombinant proteins, as well as expression vectors, and a method of imparting or enhancing production of a gymnosperm sesquiterpene synthase in a prokaryotic or eukaryotic host cell, especially in a plant, in order to impart, enhance or modify production of sesquiterpenoids, and hence to alter the flavour or fragrance of plant products, to improve disease resistance or to alter ecological interactions mediated by farnesyl diphosphate and its derivatives

XX	Sequence	2525 BP; 749 A; 498 C; 600 G; 678 T; 0 U; 0 Other;	
SQ	Query Match	32.3%; Score 723.4; DB 2; Length 2525;	
	Best Local Similarity	59.8%; Pred. No. 4.2e-214;	
	Matches 1327; Conservative	0; Mismatches 871; Indels 21; Gaps 6;	
QY	31	AATCCATCTGCATATACAGCTTGGGTGGCAAGAAATCCGTCATATGAGCGCTCTGGT	50
Db	238	ACTCCATCTGCTTATGACACAGCATGGTAGCGAGGTGCCGCCATTTGATGCTCTGCT	297
QY	91	GCACCCCAATTTCCCAAGCCTTCAATGATTTCTGACACATCAATGCCAGATGCTCG	150
Db	298	CGCCCGCAATTTCCCAACAGTTGACTGATTTTGAACCAACAGTTAAAGATGTTCA	357
QY	151	TGGGTGAGAGTGCAATTTTCTGGCGTATGACAGAGTTTTTAAACACTCTCGCCTGCC	210
Db	358	TGGGAAATTCAGTCCCACTTTCTGCTGTCGACCGTCTTCTTGCCACTCTTTCTGTGT	417
QY	211	CTCAGCTCTCAAAATATGAATAGAGGCGACATTTCAAGTGACAGAAAGGGTTGATTTG	270
Db	418	CTTGTGCTCTTAAATGGAACGTTGGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCAT	477
QY	271	AGAAACACATGGAAGAAATGAGGAGCAAGCTCACAAATCACAG- -GCCAAGTGAATTC	327
Db	478	AAGAGCAATCTGGAACTAGTAAAGAGTGAACCCGATCAAGTAGCTTGGTAAACAGCTTT	537
QY	328	GAGTCTGTTTCTCGCAATGTTAGATGAAGCAAAAGCTTTGGATTTGGATTTCTTCTTAT	387
Db	538	GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTTCCCTAC	597
QY	388	CACCTCCCTTTCATCTCCCAATCCACCAAGCGCCAGAAAGCTTCAAAAGATTTCCC	447
Db	598	GACCTGCTTATATACATCTGTCAGACTAAACGCGCAGGAAGATTTAGCAAACTTTCA	657
QY	448	CTCAATGTTCTTCAACCATCAGACGCGTGTGCTACTCTCTCGAGGGTTTCAAGAT	507
Db	658	AGGAGGAAATTTATGCGGTTCCGTGCGCATTTGTTGTTATTTCTTAGAGGGAATACA	717
QY	508	GTGGTGGACTGGCAAGAGATCACAAATCTTCAATCAAGACAGGATCAATTTTAACTCC	567
Db	718	ATAGTTGAATGGACGAATAATGGAAGTTCAAGTCAAGATGGGTCTTTCTTAAGCTCA	777
QY	568	CTGCACTACTGCTGTGCTTCTTCAATGACACACTCAAAACAAACAGATGCTCCACTTTCT	627
Db	778	CTGCTTCTACTGCTGCTTTTCAATGACACAGAGAGCGGAATGCTTGAATTTCTG	837
QY	628	AATTCGTGCTCAGCAAAATTTGGCGACTACGTTCTTGCATATACCCATTTGATCTATTT	687
Db	838	AACAGTGTGATGATCAAGTTTGGAAATTTGTTCCCTGCTGTATCTCTGTGATCTGCTG	897
QY	688	GAAACGCTCTGGCTGTGATACAGTTGACAGCTTTGGGAATCGATCGCTATTTCAAGAA	747
Db	898	GAAACGCTGTTGATCGTAGATAATATTTGACGCTTTGGAATCTATAGACATTTGAAAG	857
QY	748	GAAATCAAGAAATCTCTGATTAAGTTTATAGTACTGAGGACCGCAAGAGCGGTGGA/807	
Db	958	GAAATCAAGAAAGCTTCTGATTTATGTTTACAGGCAATTTGAA- -CGAAAGAGCAATTTGG	1014
QY	808	TGGGCAAGATGCAATCTTATCTGATGTCATGATGTCAGATGTTCTGAGAAATTTCAAGAC	867
Db	1015	TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTTCAATTTG	1074
QY	868	AGACTTCATGAGTACATGATCTTCTGATGTCATGATGTCAGATGTCGATGTCGATCTTGA	927
Db	1075	CGGCTGCATAGGTACAAATGATCTCCAGCCATTTTGAACATTTCAAGATGCAATGGG	1134
QY	928	GACTTCTTTTGTCTGCGGTCAACGCAAAATTTGTTGACCGATTAATCTTAACTTTTAT	987
Db	1135	AAATTCATTTGCTCGACCGGTCAATTTCAACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1194
QY	988	AGATGTTCAAGATATGTTTTCGGGAGAAAGATTAATGGAAGAGCTTAGACCTTCACT	1047

Db	1195	AGAGCTTCCAGCTCGCATTTCCCGAGAGAAACATTTCTTGATGAAGCTAAAGCTTCGCT	1254
QY	1048	ACAAATCATCTCAAAATGCTCTTGGCAAAACAAACGCAATTTGATAGTGGGCTGTCAAG	1107
Db	1255	ACTAAATATTTGAGAGAGCTCTTTGAGAAAGTGAGACTTCCAGTGCATGGAACAACAA	1314
QY	1108	AAGGATCTTCTCGGAGAGGTGAGTATGCTATAAAGTATCCGTTGGCATAGAGATGCGCA	1167
Db	1315	CAAACTGAGCAAGATCAATACGCGCTGAAGACTTCTTGGCATGCGATGTTCCG	1374
QY	1168	AGATTGGAGCAAGATTTACATAGAGCAATTTGGATCAATGATGCTGCTGGGGAAG	1227
Db	1375	AGAGTGGAGCAAGAGATCTCTCAAGTATTCGCCAGATTTATGACCATAGCAAA	1434
QY	1228	ACTGTGTATAAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGCCAAATTCG	1287
Db	1435	TGGTTTTACAGTACCTTACGTGAACATGAAAGTTTTTTAGAGCTGGGAAATTAGAT	1494
QY	1288	TTCAATATGTTGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGGTGAGA	1347
Db	1495	TTCAACATTTATCCAGTCCATCCCAAGAAAGAAATGAAGATTTTACAGCTGGTTTGA	1554
QY	1348	GAATCGGATTCATGATCTTACATTCACCCGCGAGCGCTGTGGAATATGTTTCTCA	1407
Db	1555	GATTCGGGGTTGCCACTATTTACCTTCGCTGGGAGAGCGCTGGAATTTCTACTTCTTA	1614
QY	1408	GTGCGGTTAGTATGTTTGGAGCCAGAAATCGCTGCTGTGAGAAATGCTATGCCAAGACT	1467
Db	1615	GTAGCGCGGGACCTATGAACCCAGTATGCCAAATGCAAGTTTCTCTTTTACAAAAGTG	1674
QY	1468	TCTTGCCTCGAGTATTTCTAGAGATCTTTTACACACACCCAGGATCTCTCGATGATCTT	1527
Db	1675	GCATGCTTGCAGACTGTTCTGAGCAGATATGATGACACTTATGGAACCTAGATGAAATG	1734
QY	1528	AAATTTGTTCTCTGAAGCGGTCGGAAGATGGGATATCTCTGTGCTGATAGGCTTCGGAT	1587
Db	1735	AAGCTATTTCTAGGCTGTGAGAAAGATGGACCTCTCTTTTACAGAAAGCTTCCAGAC	1794
QY	1588	AATCAGTTGAAATTTGCTTCTAGGCTGTGACACAGCTGAAATGGATTTTGAAAAAGAT	1647
Db	1795	TAT- -ATGAAACTATGTTACCAATCTATTATGACATAGTTTACGAGGTGCTGGGAG	1851
QY	1648	GGACTCAAGGAACAGGCGGTGATGCTGCTGGGCTATCTTCGAAAAATAGTGGAGGGCTTG	1707
Db	1852	GCAGAAAGGAACAGGCGGTGAAATTTGTCAGCTTTTTCAGAAAGGATGGAGGATTTAT	1911
QY	1708	CTCGCATCGTATACAAAGAGCCGAATGCTCGGAGCAAGATATGTCGCGACATTCAC	1767
Db	1912	CTTCTGGGTTTATGAAAGAGCTGAATGGTTAGCTGTGAGTATGTCCTTACCTTTGGAC	1971
QY	1768	GAATATGTGGAATAATGCCAAAGTGTCCATAGCTTTCGACAGTGTCTATAAATCAATC	1827
Db	1972	GAGTACATAAGAAATGGAAATCACTATTCGCGCAACGATATCTTCTGTTGATGGAGTG	2031
QY	1828	TTTTTTCAC- - -TGGAGAAATTTACTTCTGATTTACATTTTACAGCAAGTACACTTCGCTC	1884
Db	2032	TTGATAATGAGTGGCAACTCTCTTTTCGAAAGAGGCAATAGAGAAAGTAGATTATCCAG	2091
QY	1885	AAATTTCTGC- - -ATCTTGTGCTTTGACTGGACGCTAATCAATGACACCAAGACT	1938
Db	2092	AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTTCCGCGCTGGCGGATGACAGAGACA	2151
QY	1939	TACCAAGCGGAGAGAAACCGTGTGAATTTTTCAGCGGTACAGTGTCTACATGAGGAA	1998
Db	2152	TATAAGCTGAGAGGCTCGTGAGAAATTTGGCGTCCAGCATTTGAATGTTTACATGAAAG	2211
QY	1999	AATTCGGAGTGCAAGAGAAAGCTCTAAAGTCTGTTTATGTTATCGTATCATCGCAAGCA	2058
Db	2212	CATCTCTGAATGTACAGAGAGAGGCTCTCGATACATCTATAGCATTTCTGGAGCCGCG	2271
QY	2059	CTGAAGAAATTTGATTTGGGAGTTGGCCAAACCCAGAGGCAATGCCCATTTGTGTGAGA	2118
Db	2272	GTGAAGAACTGCAAGAGAGTTTCTGAAGCCG- - -ACGACGTCCTTCCGCTGCAAG	2328





Db	1255	ACTAATAATTGAGAGAGCTCTTGAGAAAGTGAGACTTCAGTGCATGGAACAACAA	1314
Qy	1108	AAGGATCTTCTGGAGAGTGAGTAGTGTATATAAGTATCCGCTGGCATAGAGATATGCCA	1167
Db	1315	CAAACCTGAGCCAAAGAGATCAATACGCGCTGAAGACTTCTTGGCATGCCAGTTCGG	1374
Qy	1168	AGATTGGAGGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG	1227
Db	1375	AGAGTGGAAAGCAAGAGATCTGCAAGTGTATCGCCACAGATTTATGCACGCATAGCAAAA	1434
Qy	1228	ACTGTGTATAAGTCTATATGTGACNACGAAATAATTTGAGCTGGCCAAATTTGGAC	1287
Db	1435	TGCGTTTACAACTACCTTACCTGAGCAATGAAGATTTTTAGAGCTGGGAAATTTAGAT	1494
Qy	1288	TTCAAATATGTCAGGCTTTACACCAAAAGAGACTCAACACATTTGTACGTGGTGAGA	1347
Db	1495	TTCAACATTTATCAGTCCATCCACCAAGAGAAATGAAGATGTTACCACTGGTTTGA	1554
Qy	1348	GAATCGGATTCAAATGATCTTAATTACCCGCGAGCGCTGTGGAAATGTTATTTCTCA	1407
Db	1555	GATTCGGGTTGCCACTATTACCTTCGCTCGGAGAGCGCTGGAAATTTCTACTCTTTA	1614
Qy	1408	GTGCGGTTAGTATGTTTGAAGCAGAAATTCGCTGTTGTAGAATTCCTATGCCAAGACT	1467
Db	1615	GTAGCGCGGGACCTATGAACCCAGTATGCCAAATGCAGGTTCCTCTTTACAAAAGTG	1674
Qy	1468	TCATTGCTCGCAGTATTCTTAGACGATCTTTACGACACCCACCGGATCTCTGGATGATCTT	1527
Db	1675	GCATGCTTCGAGACTGTTCTGGACGATATGTATGACACTTATGGAACCCCTAGATGAATTG	1734
Qy	1528	AAATTGTTCTGAAGCGTCCGAAGATGGATATCTCTGCTGGATAGCTGTTCCGGAT	1587
Db	1735	AAGCTATTCTAGAGGTGTGAGAGATGGGACCTCTCCTTTACAGAAACCTTCCAGAC	1794
Qy	1588	AATCAGTTGAAAGTTGCTTCTTAGGCTGTACCAACAGTGAATGGATTTGAAAAAGAT	1647
Db	1795	TAT---ATGAAACTATGTTACCAATCTATTATGACATAGTTTCAGAGGTGGCTGGAG	1851
Qy	1648	GGACTCAAGGAACAAAGCCGATGTGCTGGGCTATCTTCGAAAGATGGAGGGCTTG	1707
Db	1852	GCAGAGAAGGAACAGGGGCGTGAATTTGCTCAGCTTTTTCAGAAAGGGATGGAGATTAT	1911
Qy	1708	CTCGCATCGTATACCAAGAGCCGATGTCGCGCAGCAAGATGTGCCCAACATTTCAAC	1767
Db	1912	CTTCTGGGTTATTGAAGAAGCTGAATGGTTAGCTGCTGATGTATGCTTACCTACCTGGAC	1971
Qy	1768	GAATATGTGAAAAATGCCAAAGTGTCCATAGCATTGCGACAGTGTGTAACCTCAATC	1827
Db	1972	GAGTACATAAAGATGGAATCACATCTATCGCCCAACGTATATCTCTGTGAGTGGAGTG	2031
Qy	1828	TTTTTTCAC---TGGAGAAATTACTTCTGATTTACATTTTACAGCAGTAGACCTTCGGTCC	1884
Db	2032	TTGATAATGGATGGGCAACTCTCTTCGCAAGAGGCATTAGAGAAAGTAGATTATCCAGA	2091
Qy	1885	AAATTTCTGC-----ATCTGTGTCTTTGACTGACGACAAATCAATGACACCAAGACT	1938
Db	2092	AGACGTGTTCTCACAGAGCTGATAGCTCTATTTCCCGCTGGCGGATGACAGGAACA	2151
Qy	1939	TACCAGGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGGTTACGTGTACATGAGGNA	1998
Db	2152	TATAAGCTGAGAGGCTCGTGAGAAATTTGGCGTCCAGCATTTGAATGTTTACATGAAGAC	2211
Qy	1999	AATCCGGAGTCACAGAGGGAAGACTCTTAAGTCAATGTTTATGGTATCATCCACAACGCA	2058
Db	2212	CATCTGTAATGTACAGAGGAAGAGGCTCTCGATACATCTATAGCATTTGAGCCGGGG	2271
Qy	2059	CTGAAGGAATTCAAATTTGGAGTTGGCAACCCAGCGAGCAATGCCCATTTGTGTGAGA	2118
Db	2272	GTGAAGGAACCTGCAAGAGAGTTTCTGAAGCCG---ACGACGTCCCATTCGCTGCAAG	2328
Qy	2119	AGACTGCTTTCACACCTCAAGAGTGTAGCTGTTTATATGTATGACAGAGATGCTTT	2178

2329

AAGATCGTTTTCGAGGAGACAGAGTGACGATGGTGATATTCAAGGATGAGATGGAATTC

2388

2179

GGTATCTCTGACAAAGAGATGAAGACCATGTGAGCCGAACTCTTTTCGATCTCTGTGGC

2237

2389

GGTGTTCCAAATTGAAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCC

2447

RESULT 7

AA87531

AA87531 standard; DNA; 2528 BP.

AC

AA87531;

XX

08-OCT-1999 (first entry)

XX

E-alpha-bisabolene synthase computer-generated nucleic acid sequence.

XX

E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;

XX

gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour; ss.

XX

Synthetic.

XX

Abies grandis.

XX

Key

Location/Qualifiers

1. 2454

/\*tag= a

XX

WO9937139-A1.

XX

29-JUL-1999.

XX

21-JAN-1999; 99WO-US001300.

XX

22-JAN-1998; 98US-0072204P.

XX

(UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX

Croteau RB, Crock JB, Bohlman J, Jetter R, Steele CL;

XX

WPI; 1999-458569/38.

XX

P-PSDB; AAY06568.

XX

Grand fir (Abies grandis) sesquiterpene synthase genes and protein

XX

products, useful for sesquiterpenoid production.

XX

Example 11; Page 101-104; 136pp; English.

XX

This is the sequence of a computer-generated nucleic acid coding for a

XX

Grand fir (Abies grandis) E-alpha-bisabolene synthase polypeptide (see

XX

AA06568) that has conservative amino acid substitutions relative to the

XX

native E-alpha-bisabolene synthase sequence given in AAY06562. E-alpha-

XX

bisabolene synthase is a wound-inducible enzyme capable of generating

XX

multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl

XX

diphosphate. The invention provides gymnosperm sesquiterpene synthase

XX

nucleic acids and recombinant proteins, as well as expression vectors,

XX

and a method of imparting or enhancing production of a gymnosperm

XX

sesquiterpene synthase in a prokaryotic or eukaryotic host cell,

XX

especially in a plant, in order to impart, enhance or modify production

XX

of sesquiterpenoids, and hence to alter the flavour or fragrance of plant

XX

products, to improve disease resistance or to alter ecological

XX

interactions mediated by farnesyl diphosphate and its derivatives

XX

Sequence 2528 BP; 750 A; 498 C; 599 G; 681 T; 0 U; 0 Other;

XX

Sequence 2528 BP; 750 A; 498 C; 599 G; 681 T; 0 U; 0 Other;

XX

Query Match 32.3%; Score 723.4; DB 2; Length 2528;

XX

Best Local Similarity 59.8%; Pred. No. 4.2e-214;

XX

Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

XX

31 AATCCATCTGCATATGATACAGCTTGGTGGCAAGAAATCCGTCATATTGACGGCTCTGGT

90

238 ACTCCATCTGCTTATGACACAGATGGTAGCGAGGGTGGCCGCCATTGATGCTCTGCT

297

91 GCACCCCAATTTCCCCCAACGGCTTCAATGGATTTCTGAACAATCAACTGCCAGATGGCTCG

150







```
Db 1479 TGCCTTTACAGCTACCTACGTGGAACATGAAAGTTTATTAGAGCTGGGAAATTAGAT 1538
Qy 1288 TTCAATATGTGTGAGGCTTTACACCAAAAGAGAGCTCAACACATGTCTAGCTGGTGAGA 1347
Db 1539 TTCAATATTTACAGTCCATCCACCAAGAAATGAAGAAATGTATACAGCTGGTTTGA 1598
Qy 1348 GAATCGGGATTCATATCTTACATTTACATCACCAGGCGCTGTGGAAATGATTTCTCA 1407
Db 1599 GATTCCGGGTTGCCATTTATTCACCTTTCGCTCGGAGAGAGCGCTGGAATTTCTACTTCTTA 1658
Qy 1408 GTGCGGCTAGTATGTTTGAAGCAGAAATTCGCTGCTTGTAGAATTCGCTATGCAAGACT 1467
Db 1659 GTAGCGCGGGAGCTATGAACCCAGATGACCAATGCAAGTTCCTCTTTACAAAAGTG 1718
Qy 1468 TCTTGCTCGCAGTATTTCTAGACGATTTTACGACACCGACCGATCTCTGGAATGATCTT 1527
Db 1719 GCATGCTTGCAGACTGTTCTGAGCATATGTATGACACTTATGAAACCCCTAGATGAATTG 1778
Qy 1528 AATTTGTTCTCTGAAGCGTCCGAAGATGGGATATCTCTGTGCTGATAGCTGTCGGAT 1587
Db 1779 AGCTATTTCACTAGGCTGTGAGAAGATGGGACCTCTCTCTTTACAGAAACCTTCCAGAC 1838
Qy 1588 AATCAGTTGAAAGCTTTCCTAGGCTGTACAAACACAGTGAATGGATTTGAAAAGAT 1647
Db 1839 TAT-...ATGAACATATGTTACCAAAATCTATTATGACATAGTTACAGGTGGCTTGGAG 1895
Qy 1648 GAGTCAAGAAACAGCGCTGATGTCTGGGCTATCTTCGAAAAGATGAGGAGGCTTG 1707
Db 1896 GCAGAGAAGGACAGCGGCTGAATTTGGTCAAGCTTTTTCAGAAAGGATGGGAGGATAT 1955
Qy 1708 CTGCGATCGTATACCAAGAGCGGATGTCTGCGGACCAAGATGATGTGCGGACATTCAC 1767
Db 1956 CTTCTGGGTTATTTATGAAGAAGCTGAATGGTTAGTCTGCTGAGTATGTGCTTGGAC 2015
Qy 1768 GAATATGTGAAAATGCAAGGTGTCATGACCTTTCGACAGCTGCTGTAATACTCAATC 1827
Db 2016 GAGTACATAAAGATGGAATCAATCTATCGGCAAGTATATCTTCTGTTAGTGGAGTG 2075
Qy 1828 TTTTTCAC---TGGAGAATTTACTTCTGATATACATTTTACAGCAAGTAGACCTTCGGTCC 1884
Db 2076 TTGATAATGATGGGCAACTCTTTTCGCAAGAGGCATTAGAGAAAGTAGATTTATCCAGGA 2135
Qy 1885 AATTTCTGC-----ATCTTGCTCTTTGACTGGAGGACTATCAATGACCAAGACT 1938
Db 2136 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCCGCTGCGGATGACACGAAGACA 2195
Qy 1939 TACCAGCGGAGAGAAACCTGTGTAATTTGTTTCCAGCGTACAGTGCTACATGAGGAA 1998
Db 2196 TATAAAGCTGAGAGGCTGTGAGAAATGCGCTCCAGCAATGGAATGTTACATGAAGAC 2255
Qy 1999 AATCCGGAGTGCAAGAGGAAAGCTCTAAGTCTATGTTTATGGTATCATCGACAACGCA 2058
Db 2256 CATCCTGAATGTACAGAGGAAGGCTCTCGATCACATCTATAGCATTTCTGGAGCCGGCG 2315
Qy 2059 CTGAAGCAATTTAATTTGGAGTTGGCCACCCAGGAGCAATGCCCAATGTTGTGTAGA 2118
Db 2316 GTGAAGCAATGTACAAGAGAGTTTCTGAAGCCGG---ACGACGTCCTCATTTCCGCTGCAAG 2372
Qy 2119 AGACTGCTGTTTCAACACTGCAAGAGTGAATGACGCTGTTTATATGATGACAGAGATGCTTT 2178
Db 2373 AGATGCTTTTTCAGGAGACAGAGTGAATGATGATGATTTCAAGATGAGATGATTC 2432
Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGAGCGGCACTCTTTTCGATCCTGTGGC 2237
Db 2433 GGTGTTTCCAAATTTAGAAGTCAAGATCATATCAAGAGGTGCTCTCATTTGAACCGCTGCC 2491
```

RESULT 9

AAT97447

ID AAT97447 standard; cDNA; 2700 BP.

XX

AC AAT97447;

XX

```
DT 30-APR-1998 (first entry)
XX Pacific yew taxadiene synthase cDNA.
XX Pacific yew; taxadiene synthase; taxol biosynthesis; probe; immunoassay;
XX ss.
XX Taxus sp.
OS Location/Qualifiers
FH 21..2610
FT /tag= a
FT /product= "Pacific yew taxadiene synthase"
XX MO9738571-A1.
XX 23-OCT-1997.
XX 15-APR-1997; 97WO-US006320.
XX 15-APR-1996; 96US-0015993P.
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX Wildung MR, Croteau RB;
XX WPI; 1997-526123/48.
XX P-PSDB; AAW31655.
XX DNA encoding Pacific yew taxadiene synthase - used to increase taxol
XX biosynthesis.
XX Claim 10; Fig 2; 30pp; English.
XX This sequence encodes a pacific yew taxadiene synthase (TS). The sequence
XX can be used to increase taxol biosynthesis into taxoids in transformed
XX cells. The nucleic acid sequences can be modified to produce proteins
XX with improved efficiency, and to redirect the targeting of the expressed
XX polypeptide. The modifications can be to alter catalytic efficiency, to
XX modify substrate utilisation, and to alter product outcome. The
XX antibodies specific for the mature native pacific yew TS polypeptide can
XX be used to purify TS polypeptides, cloning TS homologues from Pacific yew
XX or other plant species, and as probes for blots and immunoassays
XX
SQ Sequence 2700 BP; 819 A; 541 C; 616 G; 724 T; 0 U; 0 Other;
Query Match 28.5%; Score 638.2; DB 2; Length 2700;
Best Local Similarity 57.8%; Pred. No. 1.7e-187;
Matches 1292; Conservative 0; Mismatches 903; Indels 42; Gaps 7;
Qy 1 ATGTTTTCAGTCCATGGCGGACGCTGAAACGAATCCATCTGTCATATGATACAGCTTGGGTG 60
Db 397 ATGTTTCAATGGCTCGGAGAGCGAGATATCAGTCCGCTGTCATACGACACTGCGTGGGTG 456
Qy 61 GCAGAAATTCGCTCAAT-----TGAGGCTCTGGTGCACCCCAATTTCCCAACGCTT 114
Db 457 GCAGAGCTCGCGACCAATTTCTCTGATGATCTCTGAGAAGCCAGGTTTCTCGAGCCCTC 516
Qy 115 CAATGATTTCTGAACCAATCAACTGCCAGATGGCTGCTGGGGTGCAGGATGCATTTTCTG 174
Db 517 AACTGGGTTTTCAACCAACCAAGTCCAGGATGGATCGTGGGGTATCGATCGCACTTAGT 576
Qy 175 GCGTATGACAGAGTTTTTAAACACTCTCGCTGCTCTCTCACTCTCAAAATATGGAATPAAG 234
Db 577 TTATGCGATCGATTGCTTAAACAGCAATTTCTGTTATCGCCCTCTCGGTTTGGAAACA 636
Qy 235 GCGCAATTTCAAGTGCAGAAAGGGTTGAGTTTGTGAGAAAACACATGGAAGAAATGAAG 294
Db 637 GGGCACAGCCCAAGTACAACAAGGTGCTGAGTTTATTTGACAGAAATCTAAGATTACTCAAT 696
Qy 295 GACGAAGCTGCAATCAATCACAGGCCAAAGTGGATTCGAGGTCTGTTTCTCTGCAATGTTAGAT 354
Db 697 GAGGAAGATGAGTTGTGCCCG-----GATTTCCAAATATCTTTCTCTGCTCTGCTGCAA 750
```

QY	355	GAAGCAAAAGCTTGGATTGGATCTTCTTATACCTCCCTTTCATCTCCAAATCCAC	414	QY	1435	TTCCGCTTGTAGAAATTGCCTATGCCAAGACCTTTCTTGCTCCGACGTATCTTAGACGAT	1494
DB	751	AAGGCAAAAGCGTGGGGATCAATCTTCTTACGATCTTCCATTTATCAATATTTGTGCG	810	DB	1804	TATTCGCCACTAGAAATGGCTTCCAAAATTTGGTTGTTTAAAGTCTCTTTTGTATGAT	1863
QY	415	CAAAAGCGCAGAAAAGCTTCAAAAGATTTCCCTCAATGTTCTTCAATCAACCATCAGACG	474	QY	1495	CTTTACGACACCCACCGATCTCTGGATGATCTTTAAATTTGTTCTCTGAGCGGTCGGAAGA	1554
DB	811	ACACACGGAAGCCAGGCTTACAGATGTTTC---TGGCGCAGCAGACAATATTTCCAGCC	867	DB	1864	ATGCTGACATCTTTGCAACACTAGATGAATTCAAAAGTTTCTACTGAGGAGTAAGAGA	1923
QY	475	GGGTTGCTCTACTCTCTGAGGGTTTGAAGATGTTGGTGGACTGGCAAGAGATCACAAT	534	QY	1555	TGGATATCTCTGCTGGTGGATAGGTTCCGGGATAATCAGTTTGAAGTTTCTCTTAGGG	1614
DB	868	AACATGTTGAATCGTTGAAGGCTCTCGAGGAAGTTATTGACTGGAAACAGATATGAGG	927	DB	1924	TGGATACATCTTTGCTTACATGAGATTCAGAGTGT---ATGCAAACTCTGTTTAAAGTT	1980
QY	535	CTTCAATCAAGACAGCGATCATTTTAAAGCTCCCTGCACTACTCTGTTGTTCTTCAATG	594	QY	1615	CTGTACCAACACAGTGAATGGAATTTGSAAGAGATGGAATCAAGGAACAAGCCGTTGATG	1674
DB	928	TTTCAAGTAAAGATGGAATCTTCTGAGCTCCCTGCTCCCTGCTCCACTGCTGTACTGATG	987	DB	1981	TGGTTCAAAATTAATGGAAGAAGTAATAATGATGTTGTTAAAGTACAAAGGAAGGAG	2040
QY	595	CACACTCAAAACAAACGATGCCCTTCTTCTCACTTCTGCTCAGCAAAATTTGCGGAC	654	QY	1675	CTGGGCTATCTTTCGAAAGATGAGGAGGCTTGTCTGCATCGTATACAAAGAACGCCAA	1734
DB	988	AATACAGGAGCAAAAATGTTTCACTTTCTCAACATCTGCTCGACAAATTCGGCGC	1047	DB	2041	CTCGCTCACTAAGAAACCCCTGGAGTTGTACTTCAATTTGTTATGTAACAAAGGAG	2100
QY	655	TAGTCTTCCCTTGCATACCCACTTGATCTATTGAAACGCTCTGGCTGTCGATACAGTT	714	QY	1735	TGGTCGGCAGCAAAAGTATGTCCGACATTCACAGAAATATGTGGAAATATGCCAAAGTCTC	1794
DB	1048	TGGTGCCCTGTATGTAATTCATCGATCTGCTGGAACGCTTTCGCTGGTTGATAACATT	1107	DB	2101	TGGCTTGAAGCGGATATATACCAACTTTTGAAGAGTACTTTAAAGACTTATGTCTATATCA	2160
QY	715	GAACGCTTGGGAATTCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGATACGTT	774	QY	1795	ATAGCACTTGGCAGCTGCTACTAACTCAATCTTTTCTCTGAGAGATTAATCTCTCTGAT	1854
DB	1108	GAGCATCTCGAATCGGTCGCCATTTCAAAACAAAGAAATCAAGAGAGCTCTTGATATGTC	1167	DB	2161	GTAGGCTTGGACCGGTGACCCCTACACCAATATCTACTAATGGGTGAGCTTGTGAAAGAT	2220
QY	775	TATAGGTACTGGGACCGCAAGAGGCGTGGGATGGGCAAGAGTGCATCTATTCTCGAT	834	QY	1855	TACATTTTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCAATTTCTGTCTCTTGTACTGGA	1914
DB	1168	TACAGACATTTGG---AGTGAAGGGGATCGGTTGGGCGAGACAGACGCTTGTTCAGAT	1224	DB	2221	GATGTTTGTGAGAAAGTGCATCTATCCCTCAATATATGTTTGAAGTCTTGTATCTTGAAGCTGG	2280
QY	835	GTCGATGACACTGCGATCGGTTCTAGAAATCTTGAGACTTCAATGATACATGATCTTCA	894	QY	1915	CGACTAATCAATGACACCAAGACTTACCAGGCGCAGAGAGAAACCGTGGTGAATTTGTTTCC	1974
DB	1225	CTCAACACCAAGCCCTCGGCTCGAACTCTTCGCATGACGAGATCAATGTTTCTTCA	1284	DB	2281	CGACTAACAAACGACACCAAAACATATCAGGCTGAAAGGCTCGAGGACACAACGCTCA	2340
QY	895	GATGTTCTGGAGATTTTCAGAGACGAGAAAGAGACTTCTTTTCTTGGCGGTCAAACG	954	QY	1975	AGGCTACAGTGTCTACATGAGGGGAAATCCGGAGTGACAGAGAAAGAGCTCTAAGTCAT	2034
DB	1285	GACGTTTGAATAATTTCAAGATGAAACGGCGGTTCTTCTCTGCGGGCCAAACC	1344	DB	2341	GGCATAGTAGCTATATATGAAGGATAATCCAGAGCAACTGAGAGAGATGCCATTAAGCAC	2400
QY	955	CAAAATGTTGTGACCGATAATCTTAACTTTATAGATGTTTCAAGATGTTTTCGGGA	1014	QY	2035	GTATTATGATATCATCGACCAACGACTGAGAGGAAATGAATTTGGGAGTTGGCCCAACCCAGCG	2094
DB	1345	CATGTCGAATTTGAGAAGCGTGTGAAATCTTTTCAGAGCTTCGACCTTGCAATTTCTTGAC	1404	DB	2401	ATATGTGCTGTTGTTGATCGGGCTTTGAAAGAAAGAGGCTTTGAAATATTTTCAACCATCC	2460
QY	1015	GAAAAGATAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGCC	1074	QY	2095	AGCAATGCCCATTTGTGTGAGAAGACTGCTGTTTCAACACTGCAAGAGATGAGCTGAGCTG	2154
DB	1405	GAAAGAGCTATGGAAGATCTAGAAAATTTGCAAGACCATATCTTAGAGAGCATTGCA	1464	DB	2461	AATGATATCCCATGGTTGCGAGTCTTTATTTTAACTTAGATTTGTGTGCTCAAAATC	2520
QY	1075	AAAAACAACGCAATTTGATAAGTGGGCTGTCAAGAGGATCTTCTGGAAGAGGTGAGTAT	1134	QY	2155	TTTTATATGATGACAGATGCTTTGTTGTTATCTCTGACAAAGAGATGAAAGACCATGTCAGC	2214
DB	1465	ACGAAATCTCA-----ACCAATACAAAAATATTCAAAAGAGATTTGAGTAC	1509	DB	2521	TTTTTACAAGTTTATAGATGGGTACGGAATCGCAATGCGAGATTAAGGACTATATAAGA	2580
QY	1135	GCTATAAGATATCCGTGGCATAGAGATGATGCAAGATTTGGAGGCAAGAGTTTACATAGAG	1194	QY	2215	CGAATCTTTTCGATCC	2231
DB	1510	GTGGTGAGTACCTTTGGCACATGAGTATCCACGCTTAGAAGCCAGAGTTATATTGAT	1569	DB	2581	AAAGTTTATTTGATCC	2597
QY	1195	CAATTTGGATCAATCATGCTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGC	1254	RESULT 10			
DB	1570	TCATATGACGACAAATTTATGATGGCAGAGGAAGACTCTATATAGAAATGCCATCTTTGAGT	1629	ID	AAA38931	standard; DNA; 2700 BP.	
QY	1255	AACGAAAAATTTTGGAGCTGGCCAAATTTGGATCTTCAATATGTTGGAGGCTTTACACCAA	1314	XX	AAA38931;		
DB	1630	AATTCAAAATGTTTAGAATTTGGCAAAATTTGACTTCAATATCGTACAACTTTGATCAA	1689	AC	AAA38931;		
QY	1315	AAGGAGACTCAACACATTTGTCAGCTGGTGAGAGAAATCGGATTTCAATGATCTTACATTC	1374	DT	25-AUG-2000	(first entry)	
DB	1690	GAGGATTTGAAGCTTCTTAAACAGATGTTGGAAGAAATTCGGCATGCGCAGATATAAATTT	1749	XX	Yew taxadiene synthase	DNA sequence SEQ ID NO:43.	
QY	1375	ACCGCCAGGCG					

immunological reagent; ds.  
 Taxus brevifolia.  
 WO200017327-A2.  
 30-MAR-2000.  
 17-SEP-1999; 99WO-US021419.  
 18-SEP-1999; 98US-0100993P.  
 22-APR-1999; 99US-0130628P.  
 23-AUG-1999; 99US-0150262P.  
 (KENT) UNIV KENTUCKY RES DEPT.  
 (SALK) SALK INST BIOLOGICAL STUDIES.  
 Chappell J, Manna KR, Noel JP, Starks CM;  
 WPI: 2000-292839/25.  
 P-PSDB; AAY90852.  
 Novel terpene synthase genes, useful for producing terpene hydrocarbons, e.g. fragrances or antitumor agents, are derived from known enzymes by specific amino acid alterations.  
 Disclosure; Page 408-412; 450pp; English.  
 The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 265-535 of a 348 amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha-carbon atoms (alphaC) that have interatomic distances, between each other, within tabulated ranges, have a centre point (within a sphere of radius 2.3 Angstrom) within tabulated ranges, and have an ordered arrangement of R groups (defining aa side chains), excluding specific tabulated arrangements (tables given in the specification). (I), and related enzymes, are used to produce a wide range of terpenoids (e.g. cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances, flavours, pheromones, defensive agents, pigments, antitumour agents, components of signal transduction pathways, precursors of steroid hormones and bile acids, as photoreceptors and as co-factor side chains. Some synthases with little or no catalytic activity (and nucleic acids encoding them) are used as controls in the analysis of products formed by enzymatic synthesis; as nutrient supplements; for affinity purification of isoprenoids; or to develop immunological reagents or nucleic acids for monitoring expression of terpene synthase or inheritance of the gene in plant breeding programs. The new synthases may produce novel terpene products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent sequences used in the exemplification of the present invention

Query Match 28.5%; Score 638.2; DB 3; Length 2700;  
 Best Local Similarity 57.8%; Pred. No. 1.7e-187;  
 Matches 192; Conservative 0; Mismatches 903; Indels 42; Gaps 7;  
 1 ATGTTTCAGTCCAGCGGACGGTGAACGAAATCCATCTGCATATGATACAGCTGGGTTG 50  
 397 ATGTTTCAGTCCAGCGGACGGTGAACGAAATCCATCTGCATATGATACAGCTGGGTTG 456  
 61 GCAAGAAATCCGTCAT-----TGACGCTCTGGTGACCCCAATTTCCCAAGCGTTT 114  
 457 GCGAGGCTGGCGACCAATTTCTCTGATGATCTGGAAGCCACGCGTTTCTCAGGCCCTC 316  
 115 CAATGATTTCTCAACATCACTGCGAGATGCTCGTGGGGTGAAGTGCATTTTCTG 174  
 517 AACTGGTTTTCACAAACCGCTCCAGGATGATCGTGGGGTATCGAATCGCATTTAGT 576  
 175 GCGTATGACAGAGTTTAAACACTCTCGCTCGCTCTCACTCTCAAAATATGGAAATAG 234  
 577 TTATGCGATCGATTGCTTAAACAGCAATTTCTGTTATCGCCCTCTCGGTTTGAACAA 636  
 235 GCGGACATCAAGTCGAGAAAGGGTTGAGTTTGTGAGAAACACATGGAAGAAATGAAG 294

637 GGCCACGCAAGTACACAGGTGCTGAGTTTATTGACAGAGATCTAAGATTACTCAAT 696  
 295 GACGAAGCTGCAATACACAGGCCAAGTGGATTGAGGTGCTGTTTCTGCAATGTTAGAT 354  
 697 GAGGAAGATGAGTTGTCCTCG-----GATTTCCAAATAATCTTCTGCTGCTGCTGCA 750  
 355 GAAGCAAAAGCTTGGGATGGATCTTCTCTATACACTCCCTTTCATCTCCCAATCCAC 414  
 751 AAGGCAAAAGCGTTGGGATCAATCTTCTTACGATCTTCCATTTTCAAAATATTTGTCG 810  
 415 CAAAGCGCCAGAAAAGCTTCAAAAGATTCCCTCAATGTTCTTCAATACCAATCAGACG 474  
 811 ACACACAGGGAAGCCAGGCTTACAGATGTTTCT---TGCGGACGACAGCAATATTCAGCC 867  
 475 GCGTTGCTCTACTCTCTGGAGGGTTTCAAGATGTTGGTGGACTGGCAAGAGATCACAAT 534  
 868 AACATGTTGAATGCGTTGGAAGTCTCGAGGAAGTTATTGACTGGAACAAGATTATGAGG 927  
 535 CTTCAATCAAGAGACGGATCAATTTTAAGTCCCTGCTGATCTACTGCTGTTGCTTTCATG 594  
 928 TTTCAAGTAAAGATGATCTTCTGAGCTCCCTGCTCCACTGCTGCTGCTGCTGCTGATG 987  
 595 CACACTCAAAACAAACGATGCTCCACTTCTCAACTTCTGCTGCTGCTGCTGCTGCTGCTG 654  
 988 AATACAGGGGACGAAAATGTTTCACTTTTCTCAACAATCTGCTGCAAAATTCGGCGGC 1047  
 655 TACGTTCTCTTGGCAATTTACCCACTTGTATTTTGAACGCTCTGGGCTGTCGATACAGTT 714  
 1048 TCGTGCTCTGTATGATTTCCATCGATCTGCTGGAACGCTTCTGCTGGTTGATAAAT 1107  
 715 GAACGCTTGGGAATCGATCGCTATTTCAGAGAAAGAAATCAAGAAATCTCTGGATTAGCT 774  
 1108 GAGCATCTCGGAATCGGTCGCCATTTTCAAAACAGAAATCAAGAGGAGCTCTTGATATGTC 1167  
 775 TATAGTACTGCGGACGCCGCAAGAGCGTGGGATGCGCAAGATGCAATCTTATCTCGAT 834  
 1168 TACAGACATGG---AGTGAAGGGGCTGCTGGTGGGAGAGACAGCCTTGTTCAGAT 1284  
 835 GTCGATGACACTGCCATCGGCTCTTAGAATCTTGAGACTTCTAGATACAAATGATATCTCA 894  
 1225 CTCACACACACAGCCCTCGGCTCGCAACTCTTCGCGATGACGAGATACAAATGTTCTTCA 1284  
 895 GATGTTCTGGAGAAATTCAGAGACGAGAAAGAGACTCTTCTTGGCTTGGCGGTCAACG 954  
 1285 GACGTTTTCGAAATTTTCAAGATGAAACGCGGCTTCTTCTCTCTGCGGCGCAAC 1344  
 955 CAAATGCTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGCGGA 1014  
 1345 CATGTCGAATGAGAACGCTGTTGAATCTTTTCAAGACTTTCGACCTTGCATTTCTCTGAC 1404  
 1015 GAAAGATATGGAAGAACCTTAAGACTTCACTACAAATCATCTCCAAATGCTCTTGGCC 1074  
 1405 GAAGAGCTATGACGATCTAGAAATTTGCAAGACCATATCTTAGAGAGGCACTTGCA 1464  
 1075 AAAAAACACGCAATTTGATGAGTGGCTGTCAAGAGGATCTTCTCGAGAGGTTGAGTAT 1134  
 1465 ACGAAATCTCA-----ACCAATACAAAATTTTCAAGAGATTTGAGTAC 1509  
 1135 GCTATAAGATATCCGTCGATAGAGTATGCCAAGATTGGAGGCAAGAGTATCATAGAG 1194  
 1510 GTGGTGGATGACCTTGGCAGCATGATTTCCACGCTTAGAAGCCAGAGTTATATTGAT 1569  
 1195 CAATTTGATCAATGATCTGCTGGCTGGGGAAGACTGTGTATAGATGCTATATATGAGAGC 1254  
 1570 TCATATGACGACAAATTTATGTTGACAGAGAGAGACTCTATATAGATGCAATCTTTGAGT 1629  
 1255 AACGAAAAATTTGAGCTGGCCAAATTTGAGCTTCAATATGTTGTCAGGCTTACACAA 1314  
 1630 AATTCAAAATGTTTAGAATTTGCAAAATTTGAGCTTCAATATCGTACAAATCTTTGCACTCA 1689  
 1315 AAGGAGACTCAACACATTTGCTGCTGAGAGATCGGATTCATGATCTTACATTC 1374

KW	Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;
KW	myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;
KW	terpinolene synthase; insect resistance; nutrition; ss.
OS	Abies grandis.
XX	WO200107565-A2.
PN	01-FEB-2001.
XX	24-JUL-2000; 2000WO-US020264.
PF	26-JUL-1999; 99US-00360545.
PR	(UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX	Steele CL, Bohlmann J, Croteau RB, Phillips MA;
XX	WPI; 2001-182782/18.
DR	P-PSDB; AAB69391.
XX	New nucleic acid encoding monoterpene synthases, for increasing terpene
PT	synthesis in plants, e.g. for increasing resistance to pests or for
PT	treatment of cancer.
XX	Claim 13; Page 151-154; 175pp; English.
PS	The present invention provides the protein and coding sequences of
CC	monoterpene synthases from the grand fir. These include (-)-camphene
CC	synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-
CC	limonene, (-)-alpha-pinene synthase, limonene synthase, myrcene synthase
CC	and pinene synthase. The sequences can be used to produce transgenic
CC	plants expressing high levels of the enzymes, resulting in levels which
CC	are useful in protecting against and treating cancers, and to confer
CC	insect resistance on plants
XX	Sequence 2186 BP; 722 A; 440 C; 472 G; 552 T; 0 U; 0 Other;
SQ	
Qy	Query Match 20.1%; Score 450.4; DB 4; Length 2186;
Db	Best Local Similarity 57.6%; Pred. No. 5.6e-129;
Qy	Matches 899; Conservative 0; Mismatches 631; Indels 30; Gaps 4;
Db	678 TGATCTATTGAACGCTCTCGGCTGTCGATACAGTTCGAAAGCTTGGGAATCGATCGCTA 737
Qy	390 TGATCTATTGAACGACTTTGGATGGTGCATAGCTTGAACGTTTGGGATTTGATAGACA 449
Db	738 TTTCAAGAAAGAAATCAAGAAATCTCTGGATTAACCTTTATAGTACTGGGACGCCGAAAG 797
Qy	450 TTTCAAAAAGAGATAAAATCAGCCCTTGAATATGTTTACAGTTATTTGAA 506
Db	798 AGCGTGGATGGGCAAGATGCAATCCTATTCTCATGTGATGACACTGCCATGGGTCT 857
Qy	507 AGGTATTGGATCGGTAGAGATAGTGTCTTCTCATGTCACTGCATGCTCGGGTT 566
Db	858 TAGAATCTCGAGACTTCATGGATACAAATGTATCTTCAGATGTTCTGGAGAAATTCAGAGA 917
Qy	567 TCGAACTCTTCGCTTACAGGATACAGTGTCTCTTCAGAGGTTTGAAGATTTTCAAGA 626
Db	918 CGAGAAAGAGACTTCTTTTGGTTCGCGGTCAAACGCAATTTGGTGTGACCGATAATCT 977
Qy	627 CCAAAATGGCAGTTTGGCATTTCTCTAGTAGTACAAAAGAGAGACATCAGAACCTTCT 686
Db	978 TAACTTTATAGATGTTTCAAAAGTATGTTTTCGGGAGAAAAGATTAATGGAAGACTAA 1037
Qy	687 GAATTTATATCGGGCTTCTTTTCAATTCGCTTCTTCGGGAGAAAGTATGGAAGAGCTGA 746
Db	1038 GACCTTCTACTACAAATCATCTCCAAATGCTCTTTCGCCAAAAACACGCAATTTGATAGTG 1097
Qy	747 AATTTTCTCTTCAAGATATTTGAAGAAGCCGTCGCAAAAGATTTCCCGTCT- 796
Db	1098 GGCTGCAAGAAGGATCTTCTCGGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAG 1157
Qy	797 -----CCAGTCTTTTCAAGAANAATAGACTACACTTTGGAATATGTTGGCACAC 845
Db	

b	1690	GAGGAGTTGAAGCTTCTAACAAGATGGTGGAAAGGAATCCGGCATGCGCAGATATAAATTTTC	1749
y	1375	ACCCGCCAGCGCCTGTGGAATATGTTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAA	1434
b	1750	ACTCGACACCGAGTGGCGAGGTTTATTTTTCATCAGC-----TACATTTGAACCCGAA	1803
y	1435	TTTCGCTGCTTTAGAAATTCCTATGCTCAAGAGCTTCTTGGCTCGCAGTTATTTCTAGACGAT	1494
b	1804	TATTCGCCACTAGAAATTCCTTCAAAAATTTGGTTTGTTCACAAAGTCTTTTGTATGAT	1863
y	1495	CTTTAGCACACCCAGGATCTCTGGATGATCTTAATTTGTTCTCTGAAGCGGTCCGAAGA	1554
b	1864	ATGGCTGACATCTTTGCCAACACTAGATGAATTTGAAAAGTTTCTAGGGAGTAAAGAGA	1923
y	1555	TGGGATATCTCTGTGCTGGATAGCTTTCCGGGATTAATCAGTTGAAAGTTTGTCTTCCTAGGG	1614
b	1924	TGGGATACATCTTTGCTACATGAGATTCACAGATGT---ATGCAAAATCTGCTTTAAAGTT	1980
y	1615	CTGTACAACACAGTGAATGGATTTGGAAAAGATGGACTCAAGGAACNAGGCCGTGATGTG	1674
b	1981	TGGTTCAAATTAATGGAAGAAGTAAATTAATGATGTGGTTAAGGTACAAAGACGTGACATG	2040
y	1675	CTGGGCTATCTTCGAAAAGTATGGGAGGCTTGCTCGCATCGTATACCAAGAACGCCGAA	1734
b	2041	CTCGCTCACATAAGAAAACCTCGGAGTTGTACTTCAATTTGTTATGTATACAAAGAGGGAG	2100
y	1735	TGGTCCGACGAAAGTATGTCGCGACATTCACGGAATATGTGMAATGTGMAATGCCAAGTCTCC	1794
b	2101	TGGCTTGAAGCCGGTATATACCAACTTTTGAAGAGTACTTAAAGACTTATGCTATATCA	2160
y	1795	ATAGCACTTGGACAGTCTGTAACACTCAATCTTTTCTGAGAGAAATTAATCTCTGAT	1854
b	2161	GTAGGCTTGGACGCTGTACCTCAACCAATCTACTAATGGGTGAGCTTGTGAAAGAT	2220
y	1855	TACATTTTACAGCAAGTAGACCTTCGGTCCAAAATTTCTGCATCTTGTCTTTTCACTGGA	1914
b	2221	GATGTTGTTGGAAGTGCATATCCCTCAAATATGTTTGAGCTTGTATCTCTGAGCTGG	2280
y	1915	CGACTAATCAATGACACCAAGCTTACGAGCCGAGAGAAACCGTGGTGAATGGTTTCC	1974
b	2281	CGACTAACAAACGACACCAAAACATATCAGGCTGAAAAGGCTCGAGGACAAACAGCTCA	2340
y	1975	AGCTACAGTGTCTCATGAGGAGAAAATCCGGAGTGCAACAGAGGAAGAGCTCTTAAGTCAT	2034
b	2341	GGCAATAGCTCTATATGAAGATTAATCCAGGACCACTGAGGAAGATGCCATTAAGCAC	2400
y	2035	GTATTGATATCATCGAACACGCACTGAAGAAATGAATTTGGAGTTGGCCCAACCCAGCG	2094
b	2401	ATATGTCGTGTTGTTGATCGGGCTTGAAGAAGCAAGCTTTGAATATTTCAAAACCATCC	2460
y	2095	AGCAATGCCCATTTGTGTGAGAGAGCTGCTGTTTCAACACTGCAAGAGTGAATGAGCTG	2154
b	2461	AATGATATCCCAATGGGTTGCAAGTCTTTATTTTAACTTAGATTGTGTGCTCCAAATC	2520
y	2155	TTTTATATGTACAGAGATGGCTTTGGTATCTCTGCAAGAGATGAAGACCATGTCTAGC	2214
b	2521	TTTTACAAAGTTTATAGATGGGTACGGAATCGCAATGAGGAGATTAAAGACTATATNAGA	2580
y	2215	CGAACTCTTTTCGATCC 2231	
b	2581	AAAAGTTTATTTGATCC 2597	
RESULT 11			
AAF73412			
ID AAF73412 standard; cDNA; 2186 BP.			
XX			
XX	AAF73412;		
XX	30-APR-2001 (first entry)		
DT			
XX	Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 66.		
DE			
XX			
XX			

QY 1158 AAGTATGCCAAGATTGGAGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGCTG 1217  
DB |||||  
846 AAATATGCCAAGATTGGAAACAAGGATTAATTAGATGATTTGG---ACATCCTACCAG 902  
QY 1218 GCTGGGGAGACTGTGTATAGATGCTATATGTGAGCAACGAAATAATTTGGAGCTGGC 1277  
DB |||||  
903 TCCATGGCTCAGAGAGAAAGAGCCATATCTGGACGCGAAAGCTTTTAGAATCGC 962  
QY 1278 CAAATGGACTTCAATATGGTGCAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCAG 1337  
DB |||||  
963 AAAATGGAGTTCAACATCTTCACTCCCTTCAACAGAGGAGTTACAGTATCTCTCCAG 1022  
QY 1338 CTGGTGGAGAGATCGGATTCATGATCTTACATTCACCCGCGAGGCTGTGGAAT 1397  
DB |||||  
1023 ATGGTGGATACATTCGGGTTTGGCTGAACATGACCTTTGGTCGGCATGTCACGTGGAATA 1082  
QY 1398 GTATTTCTCAGTGGCGGTGTAGTATGTTGTGACCGAGAAATTCGCTGTTGTAGAAATTCGCTA 1457  
DB |||||  
1083 CTACACCTGAGCTCTTGGCATTTGGATTTGGACTGAGCCCAACATCTTGCATCTAGATTTGGCTT 1142  
QY 1458 TGCACAGACTTTTGGCTCGCAGTTATTTAGACGATCTTTACGACACCCACGAGATCTCT 1517  
DB |||||  
1143 TGCACAAACGTGTCTATCTATCACGGTTCTGGACGATATCTACGACACTTTTCGGAACGAT 1202  
QY 1518 GGATGATCTTAATTTGTTCTGAAGCGGTTCGAGAGATGGATATCTGTGCTGGATAG 1577  
DB |||||  
1203 GGATGAATCGAACTCTTCAACGAGGCGATTTAGGAGATGGAATCCGTCGGAGAGAAACG 1262  
QY 1578 CGTTCCGGGATAATCAGTTGAAAGTTGCTTCTTAGGGCTGTCAACACAGTGAATGGATT 1637  
DB |||||  
1263 CTCCAGAAATATA---TGAAGAAATCTACATGGCACTCTACGAGCCTTAATGACAT 1319  
QY 1638 TGGAAAGATGGACTCAAGAAACAAGCCGCTGATGTGCTGGGCTATCTTTCGAAAGATG 1697  
DB |||||  
1320 GCGCGAGAGGCGAGAGAGACACAGGCGGAGACAGCTCAATTTATGCTAGAAAGGCTTG 1379  
QY 1698 GGAGGCTTGTCTGCTATGCTATACCAAGAGCGGAATGCTCGGAGCAAGATGATGGCC 1757  
DB |||||  
1380 GGAAGTTTATCTTGATTCGTATACCAAGAGCAAGATGCTGCGCAGCGGTTATCTGCC 1439  
QY 1758 GACATTCACAGATATGTGAAATGCAAGGTGCTCATAGCACTTTCGACAGTCTGACT 1817  
DB |||||  
1440 AACTTTGAGGAGTACTTAGAGAACGGAAGGTGTAGCTGTGGTCATGTCGCGCGGAT 1499  
QY 1818 AAACCTCAATCTTTTCACTGGAGAAATTAATCTCTGATTAATTTTACAGCAAGTAGACCT 1877  
DB |||||  
1500 GACACCCCTCTGACATTTGACGTAACGCTTCTGATGAGCTCTTGAAGGATAGATT 1559  
QY 1878 TCGGTCGAATTTCTGATCTTGTGCTTTGATCTGGACGACTAATCAATGACACCAAGAC 1937  
DB |||||  
1560 TCCATCGAGATTTAATGATTTGGCATCTTCTTCTTAGACTAAGAGGTGACACAGATG 1619  
QY 1938 TTACACAGGCGGAGAGAAACCGTGTGAATTTGGTTTCCAGGTACAGTGTACATGAGGA 1997  
DB |||||  
1620 CTAACAGGCGAGACGAGGAGAGAAAGCGTCAAGCATATCGTGTGTACATGAAGA 1679  
QY 1998 AAATCCGGATGACAGAGAGAAAGCTCTAAGTCTATGTTTATGCTATCATTCGACCAACGC 2057  
DB |||||  
1680 CAATCCGGATTAACAGAGGAAGATGCTCTCAATCATATCAATGCCATGATCAACGAT 1739  
QY 2058 ACTGAAGGAATGAATTTGGAGTTGGCCAAACCCAGCGAGCAATTCGCCCATTTGTGTGAG 2117  
DB |||||  
1740 AATCAAGAATTAATTTGGGAATCTTCTCAACCCGATAGCAATATTCATATGATCTGCACG 1799  
QY 2118 AGAGTCTGTTTCAACACTGCAAGATGATGAGCTGTTTATATGTACAGAGATGGCTT 2177  
DB |||||  
1800 GAAACATGCTTATGAGATACCAAGAGCTTTCCACCAACTTTACAAATATAGAGATGGCTT 1859  
QY 2178 TGGTATCTCTGACAAAGAGATGAAGACCATGTGACGCGAACTCTTTTCGATCCTGTGGC 2237  
DB |||||  
1860 CAGCGTTGCCACTCAAGAAACGAAAGTTTGGTGGAGAGAACCGTCTTGAACAGATGCC 1919

## RESULT 12

AAF73421  
ID AAF73421 standard; cDNA; 1890 BP.XX  
AC AAF73421;XX  
DT 30-APR-2001 (first entry)XX  
Grand fir monoterpene synthase coding sequence SRQ ID NO: 77.DE  
Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition; ss.XX  
OS Abies grandis.XX  
PN WO200107565-A2.XX  
PD 01-FEB-2001.XX  
PF 24-JUL-2000; 2000WO-US020264.XX  
PR 26-JUL-1999; 99US-00360545.XX  
PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.XX  
PI Steele CL, Bohlmann J, Croteau RB, Phillips MA;XX  
DR WPI; 2001-182782/18.XX  
DR P-PSDB; AAB69393.PT  
New nucleic acid encoding monoterpene synthases, for increasing terpene  
synthesis in plants, e.g. for increasing resistance to pests or for  
treatment of cancer.XX  
PS Claim 18; Page 163-165; 175pp; English.XX  
The present invention provides the protein and coding sequences of  
monoterpene synthases from the grand fir. These include (-)-camphene  
synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
and pinene synthase. The sequences can be used to produce transgenic  
plants expressing high levels of the enzymes, resulting in levels which  
are useful in protecting against and treating cancers, and to confer  
insect resistance on plantsXX  
SQ Sequence 1890 BP; 550 A; 381 C; 441 G; 518 T; 0 U; 0 Other;Query Match 18.6%; Score 416; DB 4; Length 1890;  
Best Local Similarity 57.3%; Pred. No. 2.8e-118;  
Matches 899; Conservative 0; Mismatches 620; Indels 51; Gaps 6;QY 678 TGATCTATTGTAACGCTCTGGGCTCTCGATACAGTTGAACGCTTGGGATCGATCGCTA 737  
DB |||||

366 TGATCTCATTTCAACGCTTTTAATGTCGATAACGTTGAACGTTAGGGATTGATAGACA 425

QY 738 TTTCAAGAAAGAAATCAAAGAAATCTCTGGATTACGTTATAGGTACTGGGACGCCGAAAG 797  
DB |||||

426 TTTCAAAATGAGATAAAGCAGCACTAGACTATGTTTACAGTTATTGGAA---CGAATA 482

QY 798 AGCGTGGGATGGGCAAGATGCAATCCTATTCTCGATGTCGATGACATCGCCATGGGTCT 857  
DB |||||

483 AGGCAATTGGCAGTGGAAAGTGATGTTGTTGCTGATCTCAACTCACTGCCCTGGGGTT 542

QY 858 TAGAATCTTGAGACTTCATGGATACATGATGATCTTCAGATGTTCTGGAGATTTTCAGAG- 916

543 TCGAATTTCTTCGACTACACGGATACAGTGTGTTCTTCAGATGTTGTTGGACACATTTCAAGA 602

QY 917 -----ACGAGAAAGGAGACTCTTTTGTGCTTTGGCGGTCAAACGAAATTTGGTGTGACCGA 971

603 GGNAGAGGAGAGGGGCGAGTTTGTATGTTGCGCCATCCAAACAGAGAGAGATATAAAG 662



QY	972	TAATCTTAACCTTTATAGATGTTCAAGATATGTTTCCGGAGAGAAAGATATATGGAAGA	1031
Db	663	CGTCTTGAAATTTATTCGGGCTCCCTCATTTGCTTTCCTGGGAGAAAGTTATGGAAGA	722
QY	1032	AGCTAAGACCTTCACATACAAATCATCTCCAAATGCTCTTGCCAAACCAACGCAATTTGA	1091
Db	723	GGCTGAAATCTTCTCTAAATATATTT-----AAAAGAGCCTTTACA	764
QY	1092	TAAGTGGCTGTCACAGAGATCTTCTCGAGAGTGGAGTATGCTATATAAGTATCCGTG	1151
Db	765	AAATATGCTGCTCCA---GTCTTTACAGAGATAGAGTAGCTTCTGGAGGATGGTTG	821
QY	1152	GCATAGAAGTATGCCAAGATTTGGAGCAAGAAATTCATAGAGCAATTTGGATCAATGA	1211
Db	822	GCAACAATATGCCAAGATTTGGAACAAGAACTACATCGATGTATTTGGGA-----	873
QY	1212	TGCTGGCTGGGGAAGATCTGTATAGATGCTATATGTGAGCAACGMAAAATATTGGA	1271
QY	874	-----GAGAACGATCGTGATGAGAGCTTATATATGAACATGGAGAACTTTTGA	923
QY	1272	GCTGGCCAAATTTGACCTTCAATATGTTGTCAGGCTTACACAAAGAGAGCTCAACACAT	1331
Db	924	AATTGCAAAATTTGAGTTCAATATTTTCACTCTTACAAACAGAGAGAGCTAAAGACCT	983
QY	1332	TGTCAGCTGTGGAGAGAAATCGGATTCATATGATCTTACATTCACCCGCGAGCGCTGT	1391
Db	984	CTCCAGATGTGGAAAGATTCGGGTTTCTCTACCTGACATTTTCTCGGATCTGTAAT	1043
QY	1392	GGAAATCTATTTCTCAGTGGCGTTAGTATGTTTGGAGCAAGATTCGCTGCTGTAGAAAT	1451
Db	1044	GGAAATCTCAGCTCGGATCTTGCATTTGAACTGATCGCAACATTTCCGGATTCAGACT	1103
QY	1452	TGCCTATGCCAAGACTTCTTGCCTGCGATTTATCTAGAGATCTTTACGACACCCAGG	1511
Db	1104	CGGCTTTGCCAAATTTGTCTATCTATCACGCTTTTGACGATATATACGACACTTTGG	1163
QY	1512	ATCTCTGGATGATCTTAAATTTGTTCTGAGCGCTCGAAGATGGGATATCTCTGTCT	1571
Db	1164	AACAATGGAGAGCTGGAACCTCTTCACTGCGAGATTTAAGAGATGGGATCGTCTGCCAC	1223
QY	1572	GGATAGCGTTGGGATATCAGTTGAAATTTGCTTCTCCTAGGCTGTACACACAGTGAA	1631
Db	1224	AGATTTGCTTCAGAGTATA---TGAAAGGTTGTACATGTGTGTTTACGAACCGTAA	1280
QY	1632	TGGATTTGGAAGATGACACTCAAGAACAGAGCGCTGATGTGCTGGCTATCTTCGAAA	1691
Db	1281	TGAATTTGCTCAGAGAGCAGACATCTCAAGGCCGAGAGAGCTCAAGATGCTCGAGC	1340
QY	1692	AGTATGGGAGGCTTGTCTGCAATCGTATACCAAGAAGCCGAATGGTGGCAGCAAGTA	1751
Db	1341	AGCTTTGGAGGCTATCTTGATTCGTATATGAAGAAGCTGAGTGGATCTCCAGTGGTTA	1400
QY	1752	TGTGCGGACATTCACGAATATGTGGAATAATGCCAAAGTGTCCATAGCACTTGGCAGCT	1811
Db	1401	TCGCAACGTTTGAGGAGTACATGAGACCCAGCAAGTATGTTTGGTTATTCGCATATT	1460
QY	1812	CGTACTAAATCAATCTTTTCACTGGAGAAATTTACTTCTGATTTACATTTTACAGCAAT	1871
Db	1461	CGCATTTGCAACCATCTCTCACTATGATGTTCCCTTACTCACCACATCTTCGAGGAAT	1520
QY	1872	AGACCTTCGGTCCAAATTTCTGCATCTGTGCTTTTGTGATGAGCACTAATCATGACAC	1931
Db	1521	AGACTTTCCATTCAGGTTTAAAGCTTAAATGCTTCCATCTTCGACTTTAAATAACAC	1580
QY	1932	CAAGACTTACAGGCGGAGAAACCGTGTGATTTGTTTCCAGCGTACAGTGTACAT	1991
Db	1581	TCGCTGCTACAGCGGAGAGGCGGCTTGAGAGAAAGCTTCTGTATATCGTGTATAT	1640
QY	1992	GAGGGAATTTCCGAGTGCACAGAGGAAGAGCTCTAAGTCAATGTTTATGGTATCATCGA	2051
Db	1641	GAAAGAGATCTTGGATCAACAGAGGAAGATGTCTCAATCATATCAACGCTATGTTCAA	1700
QY	2052	CAACGCACTCAAGAAATTTGAATTTGGGAGTTGGCCACCCAGCGCAATCCCCATTTGT	2111

Db	1701	TAATCTTAATCAAGAAGTGAATTTGGAGCTTCTCCGACGAGCAGCGCTCATATTGC	1760
QY	2112	TGTGAGAGACTGCTGTTTCAACACTGCAGAGCTGATGAGCTGTTTATATGTACAGAGA	2171
Db	1761	TTGCAAGAAACACGCTTTTGCATCTCTCAAGGTTCCCTTACGGCTACAAATACCGAGA	1820
QY	2172	TGCTTTGTTATCTCTGACAAAGAGATGAAGACCATGTGACGCCAACTTTTTCGATCC	2231
Db	1821	TGGTTTACGCTTGCACACAGGAACCAAGAAATTTGGGTGAGGAGAACAGTCTCTGATC	1880
QY	2232	TGTGGCGTAG	2241
Db	1881	TGTGCTTTG	1890
RESULT 13			
AAAX08663			
ID	AAAX08663	standard; cDNA; 2205 BP.	
XX	AAAX08663;		
AC	AAAX08663;		
XX	27-SEP-1999	(first entry)	
DT	27-SEP-1999		
XX	Grand Fir monoterpene synthase clone AG3.48.		
DE	Myrcene synthase; limonene synthase; pinene synthase; flavour;		
XX	monoterpene synthase; aroma; defense; plant seed; oil; meal; primer; PCR;		
KW	probe; ss.		
KW	Abies grandis.		
OS	Key	Location/Qualifiers	
XX	57.1940		
FT	/*tag= a		
FT	/product= "Monoterpene synthase"		
XX	WO9902030-A1.		
PN	21-JAN-1999.		
XX	10-JUL-1998;	98WO-US014528.	
PF	11-JUL-1997;	97US-0052249P.	
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.		
PA	Bohlmann J, Steele CL, Croteau RB;		
PI	WPI: 1999-120396/10.		
XX	P-PSDB; AAW85710.		
DR	New isolated gymnosperm monoterpene synthase DNA - obtained from Grand fir (Abies grandis), used to provide plants with modified production of monoterpenes, e.g. myrcene, limonene or pinene.		
XX	Example 3; Page 104-107; 121pp; English.		
PS	Nucleotide sequences encoding myrcene synthase, limonene synthase and pinene synthase from Grand fir may be incorporated into any organism (e.g. intact plant, animal, microbe), or derived cell culture that produces geranyl diphosphate for the production of the aforementioned enzymes or their products. The sequences when expressed in transfected cells may also be used for the production or modification of flavour and aroma properties, improvement of defense capability, and the alteration of other ecological interactions mediated by myrcene, limonene, pinene, or their derivatives. In particular they can be used for the production of plant seeds for the extraction of oil or meal		
XX	Sequence 2205 BP; 703 A; 412 C; 463 G; 627 T; 0 U; 0 Other;		
XX	Query Match	18.4%; Score 412.2; DB 2; Length 2205;	
XX	Best Local Similarity	56.8%; Pred. No. 4.7e-117;	



Matches 887; Conservative 0; Mismatches 623; Indels 51; Gaps 5;	
QY 678	TCATCTATTGTAACCGCTCTGGCTCTCGATACAGTTGAAACGCTTGGGAATCGATCGTAA 737
Db	
QY 419	TCATCTCATGCAACCGCTTTGGATAGTCGATAGCGTTGAAACGCTTGGGATAGCTAGACA 478
Db	
QY 738	TTTCAAGAAGAAATCAAGAATCTCTGATTTAGTATAGTACTGGGACCGCCAGAG 797
Db	
QY 479	TTTCAAGAAGAAATCAAGAATCTCTGATTTAGTATAGTACTGGGACCGCCAGAG 535
Db	
QY 798	AGCGTGGGATGGGCAAGATGCAATCTTCTGATGTGATGATGCACTGCGCATGGGTCT 357
Db	
QY 536	CGGCATTGATGTGGGAGAGACAGTATTGTTACTGATCTCACTCACTCGGTGGGGTT 395
Db	
QY 858	TGAATCTCTGAGACTTCATGATGATCAATGATATCTTCAGATGTTCTGGAGAAATTTGAGA 917
Db	
QY 596	TGCAACTCTTCGATTTACACGGGTACACTGTATCTCCAGAGGTTTTAAAGCTTTTCAAGA 555
Db	
QY 918	CGAGAAGGAGACTTTCTTTGCTTTGCGGTCCAAACGCAATTTGGTGTACCGGATATCT 377
Db	
QY 656	TCAAAATGGACAGTTTGTATGCTTCCCGGTCCAGACAGAGGGTGAGATCAGAAGCGTTCT 715
Db	
QY 978	TAACTTTTATAGATGTTTCAAGATATGTTTCCGGGAGAAAGATTAATGGAAGAGCTAA 1037
Db	
QY 716	TAACTTTATATCGGCTTCCCTCATTCCTTCCCTGGTGAGAAAGTTATGGAAGAGCTGA 775
Db	
QY 1038	GACCTTCACTACAAATCATCTCCAAATGCTCTTGCCAAAGAAACACGCAATTTGATAAGTG 1097
Db	
QY 776	AATCTTCTCCACAAGATATTTGAAGAAGCTCTTACAAAGATTTCCAGTCTC----- 326
Db	
QY 1098	GGCTGTCAAGAGGATCTTCTGGAGAGGTGGATGCTATTAAGATATCCGTGGGATAG 1157
Db	
QY 827	-----CGCTCTTTTCAACAGAGATAAAGTTTGTATGGAATATGGCTGGGCACAC 374
Db	
QY 1158	AAATGTCACAGATGTCAGGCAAGATTTACATAGAGCAATTTG---GATCAATCATGT 1214
Db	
QY 875	AAATTTGCCAAGATTTGAAGAGCAAGAAATTAATAGACACACTTGAAGAAGACACAGTGC 334
Db	
QY 1215	CTGGCTGGGGAAGACTGTGTATTAAGATGCTATATGTAGCAACGAAATATTTGGAGCT 1274
Db	
QY 935	ATGGCTCAATAAATGCTGGGAAGACT-----TTTAGAACT 973
Db	
QY 1275	GGCCAAATTTGAGACTTCAATATGTTGAGCGGCTTACACCAAAAGAGAGACTCAACACATTTG 1334
Db	
QY 974	TGCAAAATTTGGAGTTCAATATATTTAACTCTTACCAACAAAGGAATTAACAATATCTTTT 1033
Db	
QY 1335	CAGCTGTGTGAGAGATCGGGATTCATGATCTTACATTCACCCGCGCGGCTGTGGA 1394
Db	
QY 1034	GAGATGGTGGAAAGAGTCGGATTTGCTAAATGACATTTGCTCGGCATCGTCATGTGGA 1093
Db	
QY 1395	AATGTATTTCTCAGTGGCGGTTAGTATGTTTGAGCCAGAAATTCGCTGTGTAGAAATTCG 1454
Db	
QY 1094	ATTCTACACTTTGGCCTCTTGATTTGCCATTGACCCAAACATTCCTGCAATCAGACTAGG 1153
Db	
QY 1455	CTATGCAAGACTTTTGGCTCGCAGTTATTTAGACGATCTTTAGACACACCCACCGGATC 1514
Db	
QY 1154	CTTCGCCAAAATGTGTCTCTGTCACAGTTTGGACGATATTTAGACACACTTTTGGAAAC 1213
Db	
QY 1515	TCGTGATGATCTTAATTTGCTCTGAAAGCGTCCGAAGATGGATATCTCTGTGCTGGA 1574
Db	
QY 1214	GATTCAGGAGCTTGAATCTTTCATCTGCAATTTAAGAGATGGAATTCATCAGAGATAGA 1273
Db	
QY 1575	TAGCGTTCCGGGATAATCAGTTGAAAGTTTGTCTTCTAGGCGGTGTAACACACAGTGAATGG 1634
Db	
QY 1274	ACACCTTCCAGATATA---TGAAATGTGTGATACATGGTGGTGGTAACTGTAATGA 1330
Db	
QY 1635	ATTGGAAGAAGATGAGACTCAAGGAACAAGCCGCTGATGTGCTGGGCTATCTTCGAAAGT 1694
Db	
QY 1331	ACTGACACAGAGCGGAGAGAACTCAAGGGAGAAACACTCTCAACTATGTTCGAAAGGC 1390
Db	
QY 1695	ATGGGAGGCTTGTCTGCAATCTGTATACCAAGAGCCGATGTCGCGCAGCAAGATATCT 1754
Db	
QY 1391	TTGGGAGGCTTATTTGATTTATATATGAAGAAGCAAAATGGATCTCTAAATGTTATCT 1450
Db	

## RESULT 14

AAAF73391

ID AAF73391 standard; cDNA; 2205 BP.

XX AAF73391;

XX AAF73391;

XX 30-APR-2001 (first entry)

XX Grand fir monoterpene synthase coding sequence fragment SEQ ID NO: 31.

XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;

XX myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;

XX terpinolene synthase; insect resistance; nutrition; ss.

XX Abies grandis.

XX WO200107565-A2.

XX 01-FEB-2001.

XX 24-JUL-2000; 2000WO-US020264.

XX 26-JUL-1999; 99US-00360545.

XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.

XX Steele CL; Bohlmann J, Croteau RB, Phillips MA;

XX WPI; 2001-182782/18.

XX P-PSDB; AAB69380.

XX New nucleic acid encoding monoterpene synthases, for increasing terpene synthesis in plants, e.g. for increasing resistance to pests or for treatment of cancer.

XX

PS	Example 3; Page 135-137; 175pp; English.	
XX	The present invention provides the protein and coding sequences of monoterpen synthases from the grand fir. These include (-)-camphene synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase and pinene synthase. The sequences can be used to produce transgenic plants expressing high levels of the enzymes, resulting in levels which are useful in protecting against and treating cancers, and to confer insect resistance on plants	
XX	Sequence 2205 BP; 703 A; 412 C; 463 G; 627 T; 0 U; 0 Other;	
XX	Query Match 18.4%; Score 412.2; DB 4; Length 2205; Best Local Similarity 56.8%; Pred. No. 4.7e-117; Matches 887; Conservative 0; Mismatches 623; Indels 51; Gaps 5;	
QY	678 TGATCTATTGAAAGCGCTCTGGGCTGTCGATACAGATTGAAAGCTTGGGAATCGATCGCTA 737	QY 1515 TCTGGATGATCTTAAATTTCTCTGAAGCGGTCCGAAAGATGGGATATCTCTGTGCTGGA 1574
DB	419 TGATCTCATCAAGCGCTTTGGATAGTCGATAGCGTTGAAACGTTTGGGGATAGCTAGACA 478	DB 1214 GATTGACGAGCTTGAATCTTTCACATCTGCAATTAAGAGATGGAATTCATCAGAGATAGA 1273
QY	738 TTTCAAGAAAGAAATCAAGAAATCTCTGGATTAGCTTTATAGGTACTTGGACGCCGAAG 797	QY 1575 TAGCGTTCGGGATAATCAGTTTGAAGTTTCTCTAGGCTGTACAAACAGTGAATGG 1634
DB	479 TTTCAAGAACGAGATACATACAGCTCTGGATTATGTTTCCGTTACTGGGA---GGAAAA 535	DB 1274 ACACCTTCCAGATATA---TGAATGTGTGTACATGGTCTGTTGAAACTGTAAATGA 1330
QY	798 AGCGTGGGATGGCAAGATGCAATCTTCTGATGTCGATGACACTGCCATGGGTCT 857	QY 1635 ATTGGAAAGATGGACTCAAGGACCAAGGCCGTGTATGTCTGGGCTATCTTCGAAAGT 1694
DB	536 CGGCAATGGATGTGGGAGACAGATATTTACTGATCTCAACTCAACTGCGTTGGGGTT 595	DB 1331 ACTGACACGAGAGCGGAGAGAGACTCAAGGGAGAACACTCTCAACTATGTTCGAAAGGC 1390
QY	858 TAGAATCTGAGACTTCATGATACAAATGTATCTCAGATGTTCTGGAGAAATTTCAAGA 917	QY 1695 ATGGAGGGCTTCTCGCATCGTATACAAAGAGCGGAATGGTCGGCAGCAAGATGT 1754
DB	596 TCGAATCTTCGATTACCGGGTACATGATCTCCAGAGGTTTAAAGCTTTTCAAGA 655	DB 1391 TTGGGAGGCTTATTTTGAATTCATATATGGAAGCAAAATGGATCTCTAATGGTTATCT 1450
QY	918 CGAAGAAAGGAGACTTCTTTTCTTGGCGTCAACCGCAATTTGGTGTGACCGATATCT 977	QY 1755 GCGACATTTCAACGAATATGTGGAATGCAAGATGTCATAGCACTTGGCAGACGTCGT 1814
DB	656 TCAAAATGACAGTTTGTATGCTCCCGGTCCAGACAGGGGTGAGATCAGAAGCGTTCT 715	DB 1451 GCCAAGCTTTGAAAGAGTACCATGAGAATGGGAAGTAGCTCTGCATATCGCGTAGCAAC 1510
QY	978 TAACCTTTATAGATGTTCAAGATATGTTTCCGGGAGAAAGATATGGAAGAGCTTAA 1037	QY 1815 ACTAAACTCAATCTTTTCTACCTGGAGAAATTAATCTTCTGATTTACATTTTACAGCAAGTAGA 1874
DB	716 TAACCTTATATCGGGCTTCCCTCATTTGGAAGAGCTCTACAAAGATTCAGTCTC 826	DB 1511 ATTGCAACCCATCTCTCACTTTGAATGATGCTTCTCTGATTTACATCTTGAAGGGAATTGA 1570
QY	1038 GACCTTCACTCAATCATCTCCAAATGCTCTTGCCAAACCAACGCAATTTGATGATG 1097	QY 1875 CTTTGGTCCAAATTTCTGCACTTTGCTTCTTGACTGGAGCACTAATCAATGACACCA 1934
DB	776 AATCTTCTCCACAGATATTTGAAGAGCTCTACAAAGATTCAGTCTC 874	DB 1571 TTTTCCATCCAGGTTCAATGATTTGGCATCGTCTCTTCCGGCTACGAGGTGACACAG 1630
QY	1098 GGCTGTCAAGAAAGGATCTTCTGAGAGGTGAGTATGCTATAAGTATCGTGGCATAG 1157	QY 1935 GACTTACAGCGCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTACAGTGTCTACATGAG 1994
DB	827 -----CGCTCTTTTCAAGAGATATAAGTTTGTATGGAATATGGTGGCACAC 874	DB 1631 CTGCTCAAGGCCCATAGGGATCGTGGTGAAGAGCTTCGGTATATCATGTTATATGAA 1690
QY	1158 AAGTATGCCAAGATTTGGAGGCAAGATTTACATAGCAATTTG---GATCAATATGATGT 1214	QY 1995 GGAATAATCCGAGGTGCAAGAGGAAGCTCTAAGTCACTGTTTATGTTATCATCGACAA 2054
DB	875 AAATTTGCCAAGATTTGGAGGCAAGATTTACATAGCACACTTTGAGAAAGCACCAAGTGC 934	DB 1691 AGACAATCTCGATCAACCGAAGAAAGATGCCCTCAATCATATCAATGCCATGGTCAATGA 1750
QY	1215 CTGGCTGGGAGACTGTGTATATAGATGCTATATGTAGCAACGAAATATTTGGAGCT 1274	QY 2055 CGCACTGAAGAAATTTGAATTTGGGAGTTGGCCCAACCCAGCGAGCAATGCCCATTTCTGTGT 2114
DB	935 ATGGCTCAATAAATGCTGGGAAGAGCT-----TTTGAAGCT 973	DB 1751 CATATCAAGAAATTTAAATTTGGGAATCTTGAAGTCCCAACGACATATTTCCATCTGGC 1810
QY	1275 GGCATAATTTGACTTCAATATGGTCAGCGCTTACACCAAAAGGAGACTCAACATTTGT 1334	QY 2115 GAGAAGACTGCTGTTCAACACTGCAAGAGTGTGATGAGCTGTTTATATGTACAGAGATGG 2174
DB	974 TGCAAAATTTGGAGTTCAATATATTTAACTCTTACCAAAAGGAATTTCAATATCTTTT 1033	DB 1811 CAAGAAACATGCTTTTGACATAACAGAGACTCTCCACCATCTCTACATATATCGAGATGG 1870
QY	1335 CAGCTGGTGGAGAGATCGGATTTCAATGATCTTACATTTCACTCCCGCAGCGCTGTGGA 1394	QY 2175 CTTTGGTATCTCTGCAAGAGATGAAGACCATGTAGCCGAACTCTTTTCGATCCTGT 2234
DB	1034 GAGATGGTGGAAAGAGTCTGATTTGCTTAATTTGACATTTGCTCGCATCGTCATGTGA 1093	DB 1871 CTTTAGTGTGCAACAGGAAACAAAAAATTTGGTTATGGAACACACTCTTGAATCTAT 1930
QY	1395 AATGATTTTCTCAGTGGCGGTAGTATGTTTGGACCAAGATTTGCTGCTGTAGATTTGC 1454	
DB	1094 ATTCTACACTTTGGCGCTCTTGTATTTGCCATTTGACCAAAACATTTCTGCATTCAGATGG 1153	
QY	1455 CTATGCCAAGACTTCTTGCTGCTGAGTATTTTACAGACATCTTTACGACACCCACGGATC 1514	
DB	1154 CTTGCGCAAAATGTTGTCATCTTGTACAGTTTGGACGATATTTACGACACTTTTGGAAC 1213	

RESULT 15  
AAA38933  
ID AAA38933 standard; DNA; 1865 BP.  
XX  
AC AAA38933;  
DT  
XX 25-AUG-2000 (first entry)  
DE Grand fir delta-selinene synthase DNA sequence SEQ ID NO:47.  
XX  
XX Synthese; protein co-ordinate data; active site; modification; terpenoid;  
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
KW defensive agent; pigment; antitumour; steroid hormone;  
KW signal transduction pathway; bile acid; affinity purification;  
KW photoreceptor; enzymatic synthesis; nutrient supplement;  
KW immunological reagent; db.  
OS  
XX Abies grandis.



Thu Aug 26 08:27:53 2004

us-10-041-007-38.rng

Db	1540	GSCATCCTTAAATCGTTCCCTTGAGGAATTTAATTTGGAGCTTTATGAAGCAGGATAG---T	1596
Qy	2101	GCCCCATTGTGTGAGAGACTGCTGTTCAACACTGCAAGAGTGTATGCAGCTGTTTTAT	2160
Db	1597	GTCCCAATGTGTGCAAGAAATTCACCTTTCATATATAGGTCGAGGACTTCAATTCATCTAC	1656
Qy	2161	ATGTACAGAGATGGCTTTGGTATCTCTGACAAAGAGATGAAAGACCATGTCTAGCCGAACT	2220
Db	1657	AAATACAGAGACGGCTTATACATTTCTGACAGGAAGTAAAGGACCAGATATTCAAAAT	1716
Qy	2221	CTTTTCGATCCTGT	2234
Db	1717	CTAGTCCACCAAGT	1730

Search completed: August 24, 2004, 21:15:27  
Job time : 841.957 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	1078	48.1	2861	4	US-09-398-395A-55	Sequence 55, Appl
2	1078	48.1	2861	4	US-09-887-586A-55	Sequence 55, Appl
3	1078	48.1	2861	4	US-09-895-752-55	Sequence 55, Appl
4	1078	48.1	2861	4	US-09-903-012B-55	Sequence 55, Appl
5	1078	48.1	2861	4	US-09-900-797-55	Sequence 55, Appl
6	723.4	32.3	2424	3	US-09-234-393-1	Sequence 1, Appl
7	723.4	32.3	2424	4	US-09-360-545-15	Sequence 15, Appl
8	723.4	32.3	2424	4	US-09-865-171-1	Sequence 1, Appl
9	723.4	32.3	2424	4	US-09-398-395A-45	Sequence 45, Appl
10	723.4	32.3	2424	4	US-09-887-586A-45	Sequence 45, Appl
11	723.4	32.3	2424	4	US-09-895-752-45	Sequence 45, Appl
12	723.4	32.3	2424	4	US-09-903-012B-45	Sequence 45, Appl
13	723.4	32.3	2424	4	US-09-900-797-45	Sequence 45, Appl
14	723.4	32.3	2525	3	US-09-234-393-39	Sequence 39, Appl
15	723.4	32.3	2525	4	US-09-865-171-39	Sequence 39, Appl
16	723.4	32.3	2528	3	US-09-234-393-37	Sequence 37, Appl
17	723.4	32.3	2528	3	US-09-234-393-41	Sequence 41, Appl
18	723.4	32.3	2528	4	US-09-865-171-37	Sequence 37, Appl
19	723.4	32.3	2528	4	US-09-865-171-41	Sequence 41, Appl
20	723.4	32.3	2571	3	US-09-234-393-12	Sequence 12, Appl
21	723.4	32.3	2571	4	US-09-865-171-12	Sequence 12, Appl
22	638.2	28.5	2700	3	US-09-315-861-1	Sequence 1, Appl
23	638.2	28.5	2700	4	US-09-398-395A-43	Sequence 43, Appl
24	638.2	28.5	2700	4	US-09-887-586A-43	Sequence 43, Appl
25	638.2	28.5	2700	4	US-09-895-752-43	Sequence 43, Appl
26	638.2	28.5	2700	4	US-09-903-012B-43	Sequence 43, Appl
27	638.2	28.5	2700	4	US-09-593-253-1	Sequence 1, Appl



```

; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 55
; LENGTH: 2861
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2606)
; OTHER INFORMATION: abietadiene synthase
; US-09-887-586A-55

Query Match      48.1%; Score 1078; DB 4; Length 2861;
Best Local Similarity 58.8%; Pred. No. 0;
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

QY 1 ATGTTTCAGTCCATGGCGCAGCGTGAAACGAAATCCCATCTGCATATGATACAGCTTGGGTG 60
DB |||||
DB 378 ATGTTTAGATGATGGCTATGCGGAACGAAATCCCTCTGCATATGACACTGCTGGGTA 437
QY 61 GCAGAAATCCCGTCAATGACGGCTCTGTGCAACCCCAATTTCCCAACGCTTCAATGG 120
DB |||||
DB 438 GCAAGGATTCAGCAGCTTGATGCTCTGCAACCCCTCACTTTCTCGAGACGGTGAATGG 497
QY 121 ATTCTGAACAATCAATCGCAGATGCTCGTGGGGTGGAGTGCAATTTCTGGCGTAT 180
DB |||||
DB 498 ATTCTTCAAAATCAGTTGAAGATGGGCTTGGGGTGAAGGATTCATCTTTGGCATAT 557
QY 181 GACAGAGTTTTAAACACTCTCGCCTCGCTCCTCACTCTCAAAATATGGAATAGGGCGAC 240
DB |||||
DB 558 GACAGAAATCTGGCTACACTGATGATATTAATACCTTACCCTCTGGCGTACTGGGGAG 617
QY 241 ATTCAAGTCGAGAAAGGGTGTGATTTGTGAGAAAACACATGAGAAATGAAGGACGAA 300
DB |||||
DB 618 ACACAAAGTACAGAAAGGTATGAATCTTTCAGGACACAAGCTGGAAGATGGAAGATGAA 577
QY 301 GCTGCAATCACAGGCCAGTGGATTCGAGGTCGTCTTCTGCAATGTTAGATGAAGCA 360
DB |||||
DB 678 GCTGATAGTCATAGGCCAAGTGGATTTGAATAGTATTTCTGCAATGCTTAAAGGAAGCT 737
QY 361 AAAAGCTTGGGATTTGGATCTTCTCTATCACTCCCTCTTCACTCCCAAAATCCACCAAAAG 420
DB |||||
DB 738 AAAATCTTAGGCTTGGATCTGCTTACGATTTGCCATCTCTGAAACAAATCATCGAAAG 797
QY 421 CGCCAGAAAAGCTTCAAAAGATTCCTCAATGTTTCTTCAATCAACCATCAGACGGCGTTG 480
DB |||||
DB 798 CGGGAGGCTAAGCTTAAAGGATTCCTCACTGATGTTCTCTATGCCCTTCCAAACAAAGTTA 357
QY 481 CTCTACTCTCTGAGGGTTTGAAGATGTTGGATGTTGGATGTTGGATGTTGGATGTTGGAT 340
DB |||||
DB 858 TTGTATTTCTTGGAGGTTTACAGAAATAGTAGACTGGCGAGAAATAATGAATCAACTTCAA 917
QY 541 TCAAGAGACGGATCAATTTTAAAGCTCCCTCGCATCTACTTGTGCTTCTCATGCAACT 500
DB |||||
DB 918 TCCAGAGTGGATCAATTTCTAGCTCTCGGCATCTACAGCGGCTGTATTTCTATGCGTACA 977
QY 601 CAAAACAAACGATGCGCTTCACTTTCTCAATCTCGTCTCAGCAAAATTTGGCGACTACGTT 660
DB |||||
DB 978 GGGACAAAAGGCTTGGATTTCTTGAATCTTTGAACTTTGTAAGAAATTCGGAACCAATGTG 1037
QY 661 CCTTGCCATTAACCATTTGATCTATTTTGAACGCTCTGGGCTGTGCATACAGTTGAACGC 720
DB |||||

```

```

DB 1038 CCTTGTCACTATCCGCTTGATCTATTTTGAACGTTTGTGGCGGTTGATACAGTTGACGG 1097
QY 721 TTGGGAATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGG 780
DB |||||
DB 1098 CTAGGTATCGATCGTCAATTTCAAAGAGGAGATCAAGGAAGCATTTGGATTATGTTTACAGC 1157
QY 781 TACTGGGACGCCGAAAGAGCGGTGGATGGGCAAGATGCAATCTTATCTCTGATGTCAT 840
DB |||||
DB 1158 CATTTGGGA---CGAAAGAGGCAATTTGGATGGGCGAGAGAGAAATCTCTGTTCTCTGATTTGAT 1214
QY 841 GACACTGCCATGGGTCTTAGAATCTTAGAGATCTCATGGATACAAATGATCTTCTCAGATGTT 900
DB |||||
DB 1215 GATACAGCATAGGGCCCTTCGAATCTTGAGATTAACATGATACAAATGATCTCTCAGATGTT 1274
QY 901 CTGGAGATTTTCAGACGAGAAAGAGACTCTCTTTTGTCTTCCCGTCAAAACGCAAAAT 960
DB |||||
DB 1275 TTAATAAACATTTAGAGATGAGAAATGGGAGTTCTTTTGTCTTCTTGGTCAAAACACAGAGA 1334
QY 961 GGTGTGACCCGATATCTTAACTTTATAGATGTTTCAAGATATGTTTCCGGGAGAAAAG 1020
DB |||||
DB 1335 GGAGTTTACAGACATGTTTAAACGTCATTCGTGTTTCAATGTTTCAATTTCCCGGAGAAACG 1394
QY 1021 ATAATGGAAGAGCTAAGACCTTTCATCTACAAATCATCTCCAAATGCTCTTGGCCAAAAC 1080
DB |||||
DB 1395 ATCATGGAAGAGCAAAACTCTGTACCGAAAGGTATCTGAGGAATGCTCTGGAATATGTG 1454
QY 1081 AACGCAATTTGNTAAGTGGGCTGTCAAGAAGGATCTTCTGGAGAGGTGGATGATGCTATA 1140
DB |||||
DB 1455 GATGCTTTTGACAAATGGGCTTTTAAAGAAATATTTCCGGGAGAGGTAGATATGCACTC 1514
QY 1141 AGTATCCGTGGCATAGAAATGATGCCAAGATTTGGAGGCAAGAAAGTTTACATAGAGCAATTT 1200
DB |||||
DB 1515 AAATATCCCTGGCATAGAGTATGCCAAGGTTGGAGGCTAGAAAGCTATATTTGAAAACATAT 1574
QY 1201 GGATCAAAATGATGCTCTGGCTGGGGAAGACTGCTATAGAATGCTATATATGATGCCATGAA 1260
DB |||||
DB 1575 GGCCAGATGATGTTGGCTTGGAAAACTGATATATATGATGCCATACATTTGAAATGAA 1634
QY 1261 AAATATTTGGAGCTGGCCAAATTTGGACTTTCAATATGTTGAGGCTTTTACACCAAAAGGAG 1320
DB |||||
DB 1635 AAGTATTTAGAACTAGCGAAACTGGACTTCAATAAGGTGCGACTATACACCAAAACAGAG 1694
QY 1321 ACTCAACACATTTGCTAGCTGGTGGAGAGAAATCGGGATTCATGATCTTACATTCACCCGC 1380
DB |||||
DB 1695 CTTTCAAGATCTTCGAAAGGTGGTGAATCATCCGGTTTTCACGGATCTGAAATTTCACTCGT 1754
QY 1381 CAGCGCCTGTGGAATGATTTTCTCAGTGGGGTTAGTATGTTTTCAGCCAGAAATTCGCT 1440
DB |||||
DB 1755 GAGCGTGTGACGGAATATATTTTCTCACGGCATCCCTTATCTTTGAGCCCGAGTTTCT 1814
QY 1441 GCTTGTAGAAATTTGCCCTATGCCAAGACTTCTTGTCTGCGAGTTTATCTTAGACGATCTTTAC 1500
DB |||||
DB 1815 AAGTGCAGAGAGGTTTATACAAAACTTCCAAATTTCACTGTTATTTTAGATGATCTTTAT 1874
QY 1501 GACACCCACGATCTCTGGATGATCTTAAATGTTTCTTGAAGCGGTTCGGAAGATGGGAT 1560
DB |||||
DB 1875 GACGCCCATGATCTTTAGACGATCTTAAAGTTGTTTTCACGAATCAGTCAAAAGATGGGAT 1934
QY 1561 ATCTCTGCTCGATAGCGTTCCGGATATATCAGTTGAAAGTTTGTCTTCTAGGCGCTGTAC 1620
DB |||||
DB 1935 CTATCACTAGTGA---CCAAATGCCAACAAATGAAATATGTTTGTGGGTTTCTAC 1991
QY 1621 AACACAGTGAATGGAATTTGGAAGAGATGGACTCAAGGAAACAGGCGGTGATGTGCTGGGC 1680
DB |||||
DB 1992 AATACTTTTAAATGATATAGCAAAAGAGGACGCTGAGAGGCGGCGGATGTCTAGGC 2051
QY 1681 TATCTTGGAAAAAGTATGGAGGCGCTTGTGCGATCTGTATACCAAGAGCGGAATGGTGC 1740
DB |||||
DB 2052 TACATTTCAAAATGTTTGGAAAGTCCAACTTGAAGCTTTACACGAAAGAGCAGAAATGGTCT 2111
QY 1741 GCAGCAAAAGTATGTCGCAATTCACAGAAATATGTCGGAATATGCCAAAGTCTCATAGCA 1800
DB |||||
DB 2112 GAAGCTAAATATGTGTCATCTCTTCAATGAATACATAGAAATCGGAGTGTGTCAATAGCA 2171

```

QY	1801	CTTGGCAGAGTCTAACTCAATCTTTTCACTGGAGATTACTCTCTGATTACATT	1860	Db	378	ATGTTAGATGATGGGCTATGGCGAAACGAATCCCTCTGTCATATGACATCTGTTGGGTA	437
Db	2172	TTGGGAAACAGTGGTCTCTATTAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA	2231	QY	61	GCAAGAAATTCGTCHATTGACGGCTCTGGTCACCCCAATTTCCCAAAACGCTTCAATCG	120
QY	1861	TTACAGCAAGTAGACTCTGGTCCAAATTTCTGTCATCTTTGTCTTTTGAATCGAGACTA	1920	Db	438	GCAAGGATTCAGCAGATGATGGCTCTGACAAACCTTCACTTCTGAGACGGTTGAATGS	497
Db	2232	CTCTCCAAATTTGATCGCAATCTAGATTTCTTCAACTCATGGCTTAAACAGGGGTTTG	2291	QY	121	ATTCTGAAACAATCAACTGCCAGATGGCTCGTGGGTGAGAGTGTGATTTTCTGGGCTAT	180
QY	1921	ATCAATGACACCAAGACTTTACAGGGCCGAGAGAAACCGTGGTGTGATTTGGTTCCAGCGTA	1980	Db	498	ATTCTTCAAAATCAGTTGAAAGATGGTCTTGGGTGAAGGATTTCTACTTCTTGGCATAT	557
Db	2292	GTCATGACACCAAACTTTATCAGGCAGAGAGAGTCAAGGTGAGTGGCTTCTGCCATA	2351	QY	181	GACAGAGTTTTAAACACATCTCGCCCTGCTTCACTCTCABAATATGGAATAGGGCGAC	240
QY	1981	CAGTGTCTACATGAGGCAAAATCCGAGTGCACAGAGAAAGAGCTCTAAGTCAATGTTAT	2040	Db	558	GACAGATACTGGCTACACTTGCATGTATTTATACCCCTTACCCCTCTGGGCTACTGGGAG	617
Db	2352	CAATGTTATATGAGGACCAATCTTAANTCTCTGAGAGAGAGCTCTACAAACATGTCTAT	2411	QY	241	ATTCAAGTCAGAAAGGGGTTCAGTTTGTGAGAAAACACATGGAAGAAATGAAGGAGAA	300
QY	2041	GGTATCATCGACCAACGACTGAAGGAATGAATTGGAGTTGGCCAAACCCAGCGCAAT	2100	Db	618	ACAAAGTAGTACAGAAAGGATTTGAAATTTCTTCAAGACACAAGCTGGAAGATGAAGTAA	677
Db	2412	AGTGTCTAGGAAATGCCCTCGAAGAGTTGAATAGGGAGTTTGTGAA-----TAACAA	2465	QY	301	GCTGACAAATCACAGGCCAAGTGGATTGAGGTCTGTTTCTTCTGCAATGTTAGATGAAGCA	360
QY	2101	GCCCCATTGTTGAGAGAGCTGCTTCAACACTGCAAGAGTGTGACAGCTGTTTAT	2160	Db	678	GCTGATAGTCATAGGCCAAGTGGATTTGAAATAGTATTTCTGCAATGCTTAAAGGAGCT	737
Db	2466	ATACCGGATATTTACAAAGAGCTGGTTTTTGAACCTGCAAGATATGCAACTCTTTAT	2525	QY	361	AAAAGCTTGGGATGATCTTCCCTTATCACCTCCCTTTCATCTCCCAAAATCCCAACAAAG	420
QY	2161	ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTCAGCCGA	2217	Db	738	AAAATCTTAGGCTTGGATCTGCCCTTACGATTGCGCATCTCTGAAACAAATCATCGAAAAG	797
Db	2526	ATGCAAGGGATGGTTTGACACTATCATATGGAATTAAGAGCATGTCAAAAT	2585	QY	421	CGCAGAAAAGCTTCAAAAGATTCCCTCAATGTTCTTCAACCATCAGACGGCGTTG	480
QY	2218	ACTCTTTTCGATCTGTGGGCTAG	2241	Db	798	CGGAGGCTTAAGCTTAAAGGATTTCCCACTGATGTTCTCTATGCCCTTCCAAACAACTGA	857
Db	2586	TGCCTCTTCCAAACAGTTGCCCTAG	2609	QY	481	CTCTACTCTCTGGAGGTTTGAAGATGTCGAGCTGCGACCTGCAAGAGATCACAAAATCTTCAA	540
RESULT 3							
US-09-895-752-55							
; Sequence 55, Application US/09895752							
; Patent No. 6559297							
; GENERAL INFORMATION:							
; APPLICANT: Chappell, Joseph							
; APPLICANT: No. 65592971, Joseph P.							
; APPLICANT: Starks, Courtney M.							
; APPLICANT: Manna, Kathleen R.							
; TITLE OF INVENTION: SYNTHASES							
; FILE REFERENCE: 07678-025001							
; CURRENT APPLICATION NUMBER: US/09/895,752							
; CURRENT FILING DATE: 2001-06-29							
; PRIOR APPLICATION NUMBER: 09/398,395							
; PRIOR FILING DATE: 1998-09-17							
; PRIOR APPLICATION NUMBER: 60/100,993							
; PRIOR FILING DATE: 1998-09-16							
; PRIOR APPLICATION NUMBER: 60/130,628							
; PRIOR FILING DATE: 1999-04-22							
; PRIOR APPLICATION NUMBER: 60/150,262							
; PRIOR FILING DATE: 1999-08-23							
; NUMBER OF SEQ ID NOS: 58							
; SOFTWARE: FastSeq for Windows Version 3.0							
; SEQ ID NO 55							
; LENGTH: 2861							
; TYPE: DNA							
; ORGANISM: Abies grandis							
; FEATURE:							
; NAME/KEY: CDS							
; LOCATION: (3)...(2606)							
; OTHER INFORMATION: abietadiene synthase							
US-09-895-752-55							
Query Match 48.1%; Score 1078; DB 4; Length 2861;							
Best Local Similarity 68.8%; Pred. No. 0;							
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;							
QY	1	ATGTTTCAGTCCATGGCGAGTGAACGAATCCATCTGCATATGATACAGCTTGGGTG	60	Db	1455	GATGCCCTTTCACAAATGGGCTTTTAAAGAAATATTTCGGGGAGAGGTAGATATGCACTC	1514





Db 858 TTGTTATCTTTGGAGTTTCAAGAAATAGTAGACTGGCAGAAATAATGAACCTCAA 917  
Qy 541 TCAAGAGCGGATCATTTTAAGCTCCCTGATCTACTGCTTGCTCTTCAATGACACT 600  
Db 918 TCCAAGGATGGATCATTTCTAGCTCTCCGGCATCTACAGCGCTGATTCATGCGTACA 577  
Qy 601 CAAACAAACGATGCTCTCACTTTCTCAACTTCTGCTCTCAGCAAAATTTGGCGACTAGTT 660  
Db 978 GGGAAACAAAGTCTTGGATTTCTTGAACCTTTGCTTGAAGAAATTCGMAAACCATGTG 1037  
Qy 661 CTTTGGCAATPACCCATCTGATCTATTGAAAGCTCTGGCTCTGATCAGTTGAACGC 720  
Db 1038 CTTTGTCACTATCCGCTCTGATCTATTGAAAGCTTTGTTGGCGGTTGATACAGTTGACGG 1097  
Qy 721 TTTGGGATCGATCGCTATTCTCAAGAAAGAAATCAAGAAATCTCTGGAATACGTTTATAGG 780  
Db 1098 CTAGGTATCGATCGTCTATTCTCAAGAGGAGATCAAGAGCAATTTGGATTATGTTACAGC 1157  
Qy 781 TACTGGGACCGCGAAAGAGCGTGGGATGGCAAGATGCAATCTTATCTCTGATGTCAT 840  
Db 1158 CATTTGGGA---CGAAAGGCAATTTGGATGGCGGAGAGAAATCTGTTCTGATATTGAT 1214  
Qy 841 GACACTGCCATGGTCTTAGAATCTCTGAGACTTATCATGATCAATGATCTCTCAGATGTT 900  
Db 1215 GATACAGCCATGGGCTCTCGAATCTTGAGATTACATGATCAATGATCTCTCAGATGTT 1274  
Qy 901 CTGAGAAATTTCAAGACGAGAAAGAGACTCTTTTCTGTTGCGGCTCAAAACCAAT 560  
Db 1275 TTAACAAATTTAGAGATGAGATGGGAGTTCTTTTCTTCTTGGGCTCAACACAGAGA 1334  
Qy 961 GGTGTGACCGAATATCTTAACCTTTTAGATGTTTCAAGATGTTCTTTCGGGAGAAAG 1020  
Db 1335 GGAGTTACAGACATGTTAAACGTCATCTGTTTCAATGTTTCAATTTCCGGGAGAAACG 1394  
Qy 1021 ATATGGAAGAGCTAGACCTTCACTACAAATCAATCTCCAAATGCTCTTGGCCAAAC 1080  
Db 1395 ATCATGGAAGAGCAAAACTCTGTACCGAAGGATCTCTGGAATGCTCTGGAATAATG 1454  
Qy 1081 AACGCAATTTGATAGTGGGCTCTCAAGAAAGATCTTCTGAGAGGAGTATGCTATA 1140  
Db 1455 GATGCTTTGACAAATGGCTTTTAAAGAAATATTCGGGAGAGGTAGATGCACTC 1514  
Qy 1141 AAGTATCCGTGGCATAGAGTATGCCAAGATTTGGAGGCAAGATTTACATAGACAAATTT 1200  
Db 1515 AAATATCCCTGGCATAAGAGTATGCCAAGTTTGGAGGCTAGAGCTATATTGAAACTAT 1574  
Qy 1201 GGATCAAAATGATGCTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTAGGCAACGA 1260  
Db 1575 GGGCCAGATGATGTGGCTTGGAAAACTGTATATATGATGCCATACATTTTGAATGAA 1634  
Qy 1261 AAATATTTGGAGCTGGCCAAATTTGACTTCAATATGTTGCGAGGCTTACACCAAAAGGAG 1320  
Db 1635 AAGTATTTAGAACTAGCGAACTGGACTTCAATAGTGCATCTATACCAACCAAGAG 1694  
Qy 1321 ACTCAACATTTGTCAGCTGGTGGAGAAATCGGAATTCATGATCTTACATTCACCCGC 1380  
Db 1695 CTTTCAAGATCTTGAAGTGGTGGAAATCATCCGGTTTCAAGGATCTGAATTTCACTCT 1754  
Qy 1381 CAGCGGCTGTGGAATGATTTCTCAGTGGGCTTATGATGTTTGGAGCAAGATTCGCT 1440  
Db 1755 GAGCGGTGACCGAAATATATTCTCACCGGCACTCTTTATCTTTGAGCCCGAGTTTCT 1814  
Qy 1441 GCTTGTAGAAATTCCTATGCGAGACTTCTTGCTCGCAGTATTCTTAGACGATCTTTAC 1500  
Db 1815 AAGTGACAGAGGTTTATACAAAACTTCAATTTCACTGTTATTATTAGATGATCTTTAT 1874  
Qy 1501 GACACCAACGATCTCTGGAATGATCTTAAATGTTCTGTAAGCGGTCGGAAGATGGAT 1560  
Db 1875 GACGCCCATGATCTTTAGACGATCTTAAGTTGTTACAGAAATCAGTCAAAAGATGGAT 1934  
Qy 1561 ATCTGTGCTGATAGCGTTTCGGGATAATCAGTTGAAGTTTGTCTTCTAGGCTGTAC 1620

Db 1935 CTATCACTAGTGG---CCAAATGGCCACAAACAAATGAAATATGTTTGTGGGTTCTAC 1991  
Qy 1621 AACACAGTGAATGGATTTTGGAAAAGATGACTCAAGGAAACAAGGCCGCTGATGCTGGGC 1680  
Db 1992 AATACTTTTAATGATATAGCAAAAGAGGCGTGAAGGCAAGGGCGCATGTGCTAGGC 2051  
Qy 1681 TATCTTCGAAAAGTATGGGAGGCTTGTGCGATTCGTATATACCAAGAGCCGATGGTGC 1740  
Db 2052 TACATTCAAAATGTTTGGAAAGTCCAACCTTGAAGCTTACAGAAAGAGCAGAAATGGTCT 2111  
Qy 1741 GAGCAAAATGATGTCGACATTTCAAGAAATGTCGAAATGTCGAAATGTCGAAATGTCGAAAT 1800  
Db 2112 GAAGCTAAATATGTCGATCTCTTCAATGAATACATAGAGATGCGAGTGTCAATAGCA 2171  
Qy 1801 CTTGGCAGACTGCTATCAAACTCAATCTTTTCACTGGGAGATTTACCTCTGATTAAT 1860  
Db 2172 TTGGGAACAGTCTGTTCTCATTTAGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA 2231  
Qy 1861 TTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCACTTGTGCTTTTGAAGTGGAGCTTA 1920  
Db 2232 CTCTCCAAATTCATCGCAATCTAGATTTCTTCACTCATGGCTTAAACAGGCGTTTG 2291  
Qy 1921 ATCAATGACACCAAGACTTACAGGCGAGAGAAACCGTGTGTAATTTGGTTTCCAGCGTA 1980  
Db 2292 GTGAATGACACCAAACTTATCAGGAGAGAGGTCAGGTGAGGTGGCTTCTGCCATA 2351  
Qy 1981 CAGTGTCTACATGAGGGAATTCGGAGTGACAGAGAGAGAGCTCTTAAGTCAATGTTTAT 2040  
Db 2352 CAATGTTATATGAAGGACCATCTTAAATCTCTGAAGAGAGCTCTTACAACTGTCTAT 2411  
Qy 2041 GGTATCATGACACGACCTGAGGAAATTTGAATTTGGAGTTGGCCAAACCCAGCGAGCAAT 2100  
Db 2412 AGTGTCAATGAAAATGCCCCGAAAGATTTGAATGAGGAGTTTGTGA---TAACAA 2465  
Qy 2101 GCCCAATTTGTGTGAGAGAGACTGCTCTTCAACTGCAAGAGTATGAGCTGTTTAT 2160  
Db 2466 ATACCGATATTTACAAAGACTGTTTGAAGGAGTTTGTGA---TAACAA 2465  
Qy 2161 ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATCAAGACCATGTCAAGCGA 2217  
Db 2526 ATCAAGGGGATGTTTGAACACTATCATCATGATGGAATTTAAAGAGATGTCAGAAAT 2585  
Qy 2218 ACTTTTTCGATCTCTGTCGCTAG 2241  
Db 2586 TGCCTCTTCCAAACGTTGCTAG 2609

## RESULT 5

US-09-900-797-55  
; Sequence 55, Application US/09900797  
; Patent No. 6645762  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 66457621, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:

NAME/KEY: CDS  
LOCATION: (31...(2606)  
OTHER INFORMATION: abietadiene synthase  
US-09-900-797-55

Query Match 48.1%; Score 1078; DB 4; Length 2861;  
Best Local Similarity 58.8%; Pred. No. 0;  
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

QY 1 ATGTTTCAGTCCATGGGCGACGGTGAAACGAATCCATCTGCATATGATACAGCTCTGGGTG 50  
DB 378 ATGTTTAGATGATGGCTATGCGGAACGAATCCCTCTGCATATGACACTGCTTGGGTA 437  
QY 61 GCAGAAATTCGGTCAATGACGGCTCTGGTGCACCCCAATTTCCCAAGCGTTCAATGG 120  
DB 438 GCAGAAATTCAGACAGTTGATGGCTCTGCAACACCTTCTCTCTGAGACGGTTGAATGG 497  
QY 121 ATTCTGAACAATCAATGCGCAGATGCTCGTGGGGTGAGAGTGCATTTTCTGGCGTAT 180  
DB 498 ATTCTTCAAAATCAGTTGAAGATGGGTCTTGGGGTGAGAGTTCATCTTCTGGCATAT 557  
QY 181 GACAGAGTTTAAACACTCTCGCTCGCTCCTCACTCTCAAAATATGGAATAGGGCGAC 240  
DB 558 GACAGAAATCTGGCTACACTTGCATGATATTAATTAACCTTACCCTCTGGCGTACTGGGGAG 517  
QY 241 ATTCAAGTGCAGAAAGGGTGTGATTTGTGAGAAACACATGGAAGAAATGAAGAGCGAA 300  
DB 618 ACACAAAGTACAGAAAGGTATGAATTTCTTCAGGACACAAGCTGGAAGATGGAAGATGAA 577  
QY 301 GCTGACAAATCAGAGCCAGTGAATTCGAGGTGCTGTTCTGCAATGTTAGATGAAGCA 360  
DB 678 GCTGATAGTCATAGGCGAAGTGAATTTGAAATAGTATTTCTGCAATGCTAAAGGAAGCT 737  
QY 361 AAAAGCTTGGGATTTGGATCTTCTATCACTCCCTTTTCACTCCCAAAATCCACCAAAAG 420  
DB 738 AAAATCTTAGGTTGGATCTGCTTACGATTTGCGCATTCCTGAAACAAATCATCGAAAG 797  
QY 421 CGCCAGAAAGGCTTCAAAAGATTCCTCAATGTTTCTTCAATCAACATCAGACGGCGTTG 480  
DB 798 CGGAGGCTAAAGCTTAAAGAGATTCCTCACTGATGTTCTCTATGCCCTTCCAAACAGTTA 357  
QY 481 CTCTACTCTCGAGGTTTGAAGATGTTGGATGCTGGCAGAGATCAAAATCTTCAA 540  
DB 858 TTGATATCTTGGAGGTTTACAAGAAATAGTAGACTGGCAGAAATAATGAACACTTCAA 917  
QY 541 TCAAGAGACGGATCATTTTAAAGCTCCCTGCACTACTGCTGTTGCTTCTCATGCACT 500  
DB 918 TCCNAGATGGATCATTTCTCAGCTCTCGGCACTACAGCGGCTGTATTCATGCGTACA 977  
QY 601 CAAAACAAACGATGCTCCACTTTCTCAACTTCGTGCTCAGCAAAATTTGGCGACTACGTT 560  
DB 978 GGGAAACAAAAGTCTGGATTTCTGAACTTTGTTGTTGAAGAAATTCGGAACCAATGTG 1037  
QY 661 CCTTGCATATCCCATCTGATCTATTTGAAGCTCTGGCTGTGCATACAGTTGAAGCG 720  
DB 1038 CCTTGTGCACTATCCGCTGATCTATTTGAACGTTTGTGGCGGTGATACAGTTGAGCGG 1097  
QY 721 TTGGGAATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGATTAAGTTATAGG 780  
DB 1098 CTAGGTATCGATCGTCTTTTCAAGAGAGATCAAGAGCATTTGATATGTTTACAGC 1157  
QY 781 TACTGGGACCGCGAAAGAGCGTGGGATGGGCAAGATGCAATCTTATCTCGATGTCGAT 340  
DB 1158 CATTTGGGA---CGAAAGAGCATTTGGATGGCGGAGAGAAATCTGTTCTCGATATTTGAT 1214  
QY 841 GACATCTGCCATGGGCTTTAGATCTTGAAGCTTTCATGATGATGATATCTCAGATGTT 900  
DB 1215 GATACAGCCATGGGCGCTTCAATCTTTGAGATTAATGATGATGATGATGATGATGATGAT 1274  
QY 901 CTGGAGATTTTCAAGACGAGAAAGAGACTCTTTTGTCTTTCGCGGTCAAAACGCAAAAT 960  
DB 1275 TTAAACAAATTTAGATGAGATGGGAGTCTTTTGTCTTTCTTGGGTCAACACAGAGA 1334

QY 961 GGTGTGACCGGATTAATCTTAACCTTTTATAGATGTTTCAAGATATGTTTTCGGGAGAAAAG 1020  
DB 1335 GGAGTTTACAGACATGTTAAACGTCATCTGTTTCAATCTGTTTCAATCTTTCGGGAGAAAAG 1394  
QY 1021 ATAATGGAAGAGCTAAGACCTTTCACCTACAAAATCATCTCCAAAATGCTCTTTCGCAAAAAC 1080  
DB 1395 ATCATGGAAGAGCAAAACTCTGTACGGAAGGTATCTGAGGAATGCTCTGGAANAATGTG 1454  
QY 1081 AACGCAATTTGATAAGTGGGCTGTCAAGAAGGATCTTCTCGGAGAGGTGGAGTATGCTATTA 1140  
DB 1455 GATGCTTTTGACAAATGGGCTTTTAAAAAGAAATATTCGGGGAGAGGTAGAGTATGCACTC 1514  
QY 1141 AAGTATCCGTGGCATAGAGATATGCCAAGATTCGAGGCAAGAAGTTTACATAGAGCAATTT 1200  
DB 1515 AAATATCCCTGGCATAGAGTATGCCAAGGTTCGAGGCTTGAAGCTATATTTGAANAATAT 1574  
QY 1201 GGATCAAAATGATCTCTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTGAGCAACGAA 1260  
DB 1575 GGCCAGATGATGTTGGCTTGGAAANAATGTATATATATGATGCCATACATTTGAAATGAA 1634  
QY 1261 AAATATTTGGAGCTGGCCAAATTTGCAATTTGCTGAGGCTTTTACCAAAAAGGAG 1320  
DB 1635 AAGTATTTAGAACTAGCGAAACTGGACTTCAATAAGGTGCACTTATACACCAAAACAGAG 1694  
QY 1321 ACTCAACACATTTGCTGAGCTGGTGGAGAGAAATCGGATTTCAATGATCTTATCATTCACCGC 1380  
DB 1695 CTTTCAAGATCTTTCGAAGGTGGTGAANAATCATCCGGTTTTCACGGATCTGAAATTTTCACTCGT 1754  
QY 1381 CAGCGCCTGTGGAATGATTTTCTCAGTGGCGGTTAGTATGTTTTCAGCCAGCAATTCGCT 1440  
DB 1755 GAGCGGTGACGGAANAATATATTTCTCACCGGCACTCTTTATCTTGTGAGCCGAGTTTCT 1814  
QY 1441 GCTTGTAGAAATTTGCTTATGCAAGACTTCTTGGCTCGCAGTATTTCTAGACGATCTTTTAC 1500  
DB 1815 AAGTGCAGAGAGGTTTATACAAAACCTTCCAAATTCACCTGTTATTTTGTAGATGATCTTTAT 1874  
QY 1501 GACACCAACGATCTCTGGAGTATCTTAAATGTTTCTCGAAGCGGTTCGGAAGTGGAT 1560  
DB 1875 GACGCCATGGATCTTTTAGACGATCTTAAAGTTGTTTTCACAGAAATCAGTCAAAAGATGGAT 1934  
QY 1561 ATCTGTGCTGGATAGCGTTGGGATATACGTGGAAGCTTGTCTTCTTAGGCTGTAC 1620  
DB 1935 CTATCTAGTGA---CMAATGCCAACAAATGAAATATGTTTGTGGGTTCTTAC 1991  
QY 1621 AACACAGTGAATGGATTTGGAAAGATGGAATCAAGGAAACAAAGCCGCTGATGCTGGGC 1680  
DB 1992 AATACCTTTTAAATGATATAGCAAAAGAGAGCGTGAGAGGCAAGGCGCGATGCTAGGC 2051  
QY 1681 TATCTTCGAAAAGTATGGAGGCTTGTGCGATCGTATACCAAGAGAGCGGAATGGTCG 1740  
DB 2052 TACATTTCAAAATGTTTGGAAAGTCCAACTTTGAAGCTTTACACGAAAGAGCAAGATGGTCT 2111  
QY 1741 GCAGCAAGATGTTGGCGGACATTTCAACGAATATGTGMAAATGCCAAAGTGTCCATAGCA 1800  
DB 2112 GAAGCTAAATATGTCGCCATCTTCAATGAATACATAGAAATGCGAGTGTCTCAATAGCA 2171  
QY 1801 CTTTGCACAGTCTGATCTAAACTCAATCTTTTTCACCTGGAGAAATTAATCTTCTGATTAAT 1860  
DB 2172 TTGGGAACAGTCTGTTCTCATAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA 2231  
QY 1861 TTACAGCAAGTATAGCTTGGTCCAAATTTTCTGATCTTGTGCTTTTGAATGCACTGAGCACTA 1920  
DB 2232 CTCTCAAAATTTGATCGCAATCTAGATTTCTTCAACTCATGGGCTTAAACAGGCGGTTG 2291  
QY 1921 ATCAATGACCAAGACTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
DB 2292 GTGAATGACCAACCAAACTTATCAGGAGAGAGAGGTCAGAGTGGGTTCTTCTGCCATA 2351  
QY 1981 CAGTGTCTACATGAGGGAATAATCCGAGTGCACAGAGAGAAAGCTCTTAAGTCTATGTTTAT 2040  
DB 2352 CAATGTTATATGAGAGCACTCTTAAATCTCTGAGAGAGAAAGCTCTTACAACTGCTAT 2411  
QY 2041 GGTATCATCGACAAACGCACTGAAGAGAAATTTGAATTTGGAGTTTGCCCAACCCAGGAGCAAT 2100

Db	2412	AGTGTCTGGAATGCGCTCGAAGAGTTGAATAGGAGTTGTGAA-----TAACAAA	2465	Qy	448	CTCAATGTTCTTTCAAAACCATCAGACGGGTTGCTCTACTCTCTGGAGGGTTTGGCAAGAT	507
Qy	2101	GCCCAATGTTGTGAGAGAGCTGCTGTTCAACACTGCAAGAGTGATGACAGCTGTTTTAT	2160	Db	554	AGGGAGGAATTTATGCGGTTCCGTCGCCATTTGTTGTAATCTTTAGAGGGAATACAAGAT	613
Db	2466	ATACCGGATATTACAAAAGACTGGTTTTTGAACCTGCAAGATATGCAACTCTTTAT	2525	Qy	508	GTGTTGGAATGCGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTTAAAGCTCC	567
Qy	2161	ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTCAGCCGA	2217	Db	614	ATAGTTGAATGCGAAACGAATAATGGAAGTTCAAAGTCAGGATGGGTCTTTCTTAAGCTCA	673
Db	2526	ATGCAAGGGATGGTTTGACACTATCACATGATATGGAATTAAGAGCATGTCAAAAT	2585	Qy	568	CCTGCATCTACTGCTGTCTTCTTCATGCACACTCAAAACAAAGAGATGCGCTCCACTTTCTC	627
Qy	2218	ACTCTTTTCGATCCTGTGGCGTAG	2241	Db	674	CCTGCTTCTACTGCTGCGTTTTCATGCACACAGGAGACGCGAAATGCGCTTGAATCTTG	733
Db	2586	TGCTCTTCCAAACCAAGTTGCCTAG	2609	Qy	628	AACTTCGTCTCAGCAAAATTTGGCGACTACGTTCTTGCCTATCCACTTGATCTATTT	687
RESULT 6				793	Db	734	AAACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGTATCTCTGTGATCTGCTG
US-09-234-393-1				747	Qy	688	GAACGCTCTGGGCTCTCGATACAGTTTGAACGCTTCGGGAATCGATCGCTATTTTCAAGAAA
; Sequence 1, Application US/09234393A				853	Db	794	GAACGCTCTGTGATCGTAGATAATATTTGACGCTTGGAAATCTATAGACACTTTTGAAGA
; Patent No. 6265639				807	Qy	748	GAATCAAGAATCTCTGGATTACGTTTATAGTACTTGGGACGCCGAAAGAGCGCGTGGGA
; GENERAL INFORMATION:				910	Db	854	GAATCAAGGAGCTCTTGATTATGTTTACAGGCAATTGGAA---CGAAAGAGGAATTTGG
; APPLICANT: Croteau, Rodney B				867	Qy	808	TGGCAAGATGCAATCTTCTGATGTCGATGACACTGCCATGGTCTTTAGAACTCCTG
; APPLICANT: Bohman, Jorg				970	Db	911	TGGGCGACACTAAATCCCATAGCAGATCTTGGAGACCACTGCTTTGGGATTTTCAATGCTT
; APPLICANT: Steele, Christopher L				927	Qy	868	AGACTTCATCGATCAATGATCTTTCAGATGTTCTGGAGATTTTCAGAGACGAGAAAGGA
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)				1030	Db	971	CGGCTGCATAGTACAATGATCTCCAGCCATTTTTCAGCAACTTCAAAGATGCCAATGGG
; TITLE OF INVENTION: AND METHODS OF USE				987	Qy	928	GACTTCTTTTGTCTTGGCGGTCAAAACGCAAAATTTGGTGTGACCGATATCTTAACCTTTAT
; FILE REFERENCE: WSUR113345				1090	Db	1031	AAATTCATTTGCTCGACCGGTCAATTCACAAGAGATGTAGCAAGCATGCTGAAATCTTTAT
; CURRENT APPLICATION NUMBER: US/09/234.393A				1047	Qy	988	AGATGTTTCAAGATGATGTTTTCGGGAGAAAGATATGGAAGAGCTTAAGAGAGCTTCACT
; CURRENT FILING DATE: 1999-01-20				1150	Db	1091	AGAGCTTCCAGCTCGCATTTTCCGGAGAAACATTTCTTGATGAGCTTAAAGCTTCGCT
; EARLIER APPLICATION NUMBER: 60/072.204				1270	Qy	1048	ACAAATCATCTCCAAATGCTCTTTGCCAAAACCAACGATTTGATTAAGTGGGCTGTCAAG
; EARLIER FILING DATE: 1998-01-22				1227	Db	1151	ACTAAATTTTGGAGAGAGCTCTTTGAGAAAGTGGAGACTTCCAGTGCATGGAACAACAA
; NUMBER OF SEQ ID NOS: 55				1330	Qy	1108	AAGGATCTTCTGAGAGAGTGATGCTATAAGTATCCGTGCGCATAGAAATGATGCGCA
; SOFTWARE: PatentIn Ver. 2.0				1287	Db	1211	CAAAACCTTGAGCAAGAGATCAAAATACGCGCTGAAGACTTCTTGCGATGCCAGTGTCCG
; SEQ ID NO 1				1390	Qy	1168	AGATTGGAGGCAAGAAAGTTTACATAGAGCAATTTGGATCAAAATGATGTCTGCTGGGGAAG
; LENGTH: 2424				1347	Db	1271	AGATTGGAGGCAAGAGATCTGTCAAGTGTATCGCCAGATTTGACCGCATAGCAAAA
; TYPE: DNA				1450	Qy	1228	ACTGTGTATAAGATGCTATATGTAGAGCAACGAAATAATTTTGGAGCTGCCAAATTTGAC
; ORGANISM: Abies grandis				1407	Db	1331	TGCGTTTACAAGCTACCTTACGTTGAACAAATGAAAGTTTATAGAGCTGGGAAATTTAGAT
US-09-234-393-1				1510	Qy	1288	TTCAATAATGGTGCAGGCTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGGTGGAGA
Query Match 32.3%; Score 723.4; DB 3; Length 2424;				1467	Db	1391	TTCAACATTTATCCAGTCCATCCCAAGAGAAATGAAAGATGTTTACAGCTGGTTTGA
Best Local Similarity 59.8%; Pred. No. 3.8e-222;				1527	Qy	1348	GAATCGGGAATTCATATGATCTTACATTTCAACCGCAGCGGCTGTGGAATATGATTTCTCA
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;				1570	Db	1451	GATTCGGGTTGCCACTATTACCTTCGCTCGGAGAGGCGCGCTGGAATTTCTACTTTT
Qy	31	AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCGTCAAATTCAGCGCTCTGCT	50	Qy	1408	GTGGCGGTTTGTATGTTTGGAGCCAGAAATTCGCTGCTGTGTAGAAATTTGCCATGCCAAGACT	
Db	134	ACTCCATCTGCTTATGACACAGCATGGGTAGCGGGTGCCCGCCATTTGATGGCTCTGCT	193	Qy	1511	GTAGCGCGGGGACCTTATGAACCCAGTATGCCAAATGAGGTTTCTCTTTTCAAAAGTG	
Qy	91	GCACCCCAATTTCCCAACGCTTCAATGGATTCTGAACAAATCAACTGCCAGATGGCTCG	150	Qy	1468	TCTTGTCTCGAGTATTTCTTAGACGATCTTTTACGACACCCACCGATCTCTGGATGATCTT	
Db	194	CGCCCGCAATTTCCCAACAGTTGACTGGATTGTAAGAAATCAACAGTTAAAGATGTTCA	253	Qy	1571	GCAATGCTTGCAGACTGTTCTGACGATATGATGACACTTATGGAACCCCTAGATGAATTG	
Qy	151	TGGGGTGAAGAGTGATATTTCTGCGGTATGACAGATTTTAAACACTCTGCGCTGCGCTC	210	Qy	1528	AAATTTGTTCTCTGAAAGGGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGTTCGGGAT	
Db	254	TGGGGAATTCAGTCCCACTTCTGCTGTCGCGACCGCTCTCTTGCCACTCTTCTTGTTG	313				
Qy	211	CTCACTCTCAAAATATGAAATGAAGGCGACATTCAGTGCAGAAAGGGGTTCAGTTTGTG	270				
Db	314	CTTGTGCTCTTAAATGGAACGTTTGGGGATCTGCAAGTAGAGCAGGGAATTTGAATTCATA	373				
Qy	271	AGAAACACATGGAAGAAATGAAGGAGCAAGCTGACAAATCAAG---GCCAAGTGGATTTC	327				
Db	374	AGAGCAATCTGGAATAGTAAAGATGAACCGATCAAGATAGCTTGTGTAACAGACTTT	433				
Qy	328	GAGGTCGTGTTTCTCGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATCTTCTCTAT	387				
Db	434	GAGATCATATTTCTTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC	493				
Qy	388	CACCTCCCTTTCATCTCCCAATCCACCAAGCGCCAGAAAAAGCTTCAAAAGATTTCCC	447				
Db	494	GACCTGCTTATATACATCTGTTGCACTTAACGGCAGGAAGATTAGCAAACTTCA	553				

1631 AAGCTATTCAGTGGGCTGTGAGAGATGGAGCTCTCTTTTACAGAAAACCTTCCAGAC 1690  
1588 AATCAGTTGAAGTTTCTCTAGGGCTGTACAAACAGTGAATGGATTGGAAAAGAT 1647  
1691 TAT---ATGAACATCTATTACCAAAATCTATTATGACATAGTTCCAGAGGTGGCTGGAG 1747  
1648 GGAATCAAGAACAGAGCCCTGATGCTGGCTATCTTCGAAAAGTATGGAGGGCTTG 1707  
1748 GCAGAGAAGAACAGGGGCTGAAATGGTCAAGCTTTTTCGAAAAGGATGGAGGATTAT 1807  
1708 CTCGCATCGTATACCAAGAAGCCGAATGGTTCGCGACAAAGTATGTGCCGACATTCAC 1767  
1808 CTTCTGGGTTATATGAAGAGCTGAATGGTATGCTGCTGATGATGCTCTACCTGGAC 1867  
1768 GAATATGTGGAATGCCAAAGTGTCCATAGCACTTTCGCGACGTCTGTAATACTCAATC 1827  
1868 GAGTACATAAAGAATGGAATCACATCTATCGGCCAACGTATCTCTGTTGAGTGGAGTG 1927  
1828 TTTTTCAC---TGGAGAAATCTCTCTGATATACATTTTACAGCAAGTAGACCTTCGGTCC 1884  
1928 TTGATATGATGGGCAACTCTTTTCGCAAGAGGCATTAGAGAAAAGTAGATTATCCAGGA 1987  
1885 AATTTCTGC-----ATCTGTGTCTTTCGACTGGACGACTAATCAATGACACCAAGCT 1938  
1988 AGAGTGTTCTCAGAGCTGGAATAGCTCATTTTCCGCTTGGCGGATGACAGAGACA 2047  
1939 TACCAGCCGAGAGAAACCGTGTGTAATGGTTTTCAGCGTACAGTGTCTACATGAGGGAA 1998  
2048 TATAAAGCTGAGAAGGCTCGTGGAGAAATGCGTCCAGCAATGAATGTTACATGAAAGAC 2107  
1999 AATCCGGAGTGCACAGAGGAAGCTCTAAGTCAATGTTTATGGTATCATCGACAGCA 2058  
2108 CATCTCTGAATGTACAGAGGAAGGCTCTCGATCACAATCTATAGCAATTCGAGCCGGCG 2167  
2059 CTCAGGAATGTAATGGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2118  
2168 GTGAAGNACTGACAGAGAGTTTCTGAAGCCCG---ACGACGTCCCATTTGGCTGCAAG 2224  
2119 AGACTGTCTTCAACCTGCAAGAGTGTAGTGTGTTTATATGTATGACAGAGATGGCTTT 2178  
2225 AAGATGCTTTTCGAGGAGACAGAGTGTGATATTCAGGATGGAGTGGATTTC 2284  
2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGACGCCGAATCTTTTCGATCTCTGTGGC 2237  
2285 GGTGTTTCCAAATTAGAAGTCAAAGATCATATCAAAGAGTGTCTCATTTGAACCGCTGCC 2343

## RESULT 7

US-09-360-545-15  
; Sequence 15, Application US/09360545  
; Patent No. 6429014  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B  
; APPLICANT: Bohlmann, Jorg  
; APPLICANT: Steele, Christopher L  
; APPLICANT: Phillips, Michael A  
; TITLE OF INVENTION: MONOTERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)  
; FILE REFERENCE: waur13885  
; CURRENT APPLICATION NUMBER: US/09/360,545  
; CURRENT FILING DATE: 1999-07-26  
; EARLIER APPLICATION NUMBER: 60/052,249  
; EARLIER FILING DATE: 1997-11-07  
; EARLIER APPLICATION NUMBER: PCT/US98/14528  
; EARLIER FILING DATE: 1998-07-10  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (2) ..(2350)  
; OTHER INFORMATION: Clone AG1.28  
US-09-360-545-15  
Query Match 32.3%; Score 723.4; DB 4; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 3.8e-222;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;  
QY 31 AATCCATCTGCATATATACAGCTTGGGTGGGCAAGAATTCCTGCAATGTACGGCTCTGGT 90  
DB 134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGAGGTGCCGCCATTGATGGCTCTGCT 193  
QY 91 GCACCCCAATTTCCCAACAGCTTCAATGGAATCTGAACAACTCAACTGCCAGATGGCTCG 150  
DB 194 CGCCCGCAATTTCCCAACAGCTTGAATTTGAAAACCAAGTTTAAAGATGGTTCA 253  
QY 151 TGGGGTGAGGAGTGCATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTCGCTC 210  
DB 254 TGGGGAAATTCAGTCCCACTTTCTGCTGTCGACCGTCTTCTGCCACTCTTTCTTGTT 313  
QY 211 CTCACCTCTCAAAATATGGAATTAAGGCGCAATTCAGTGCAGAAAAGGGTGTGATTTGTG 270  
DB 314 CTTGTGCTCTTAAATGGAACGTTTGGGATCTGCAAGTAGAGCAGGAAATGAATTCATA 373  
QY 271 AGAAAACATGGAAGAAATGAAGGACGAAGCTGACAACTACACAG---GCCAAGTGGATTC 327  
DB 374 AGAGCAATCTGGAACTAGTAAAGGATGAACCGTCAAGATAGCTTGGTAAACAGACTTT 433  
QY 328 GAGGTGCTGTTTCTCTCAATGTTAGATGAAGCAAAAAGCTTGGATTTGGATCTTCTCTAT 387  
DB 434 GAGTCAATATTTCTTCTGTTTAAAGAAAGCTCAATCTCTGCGCTCGACATTTCCCTAC 493  
QY 388 CACCTCCCTTTCTCCAAATCCCAAAAGCGCCAGAAAAGCTTCAAAAAGATTTCC 447  
DB 494 GACCTGCTTATATACATCTGTTGACACTAAACCGCAGGAAAGATTAGCAAAACTTCA 553  
QY 448 CTCATGTTCTTCAATACCATCAGACGGGTGCTCTACTCTCTGAGGGSTTTCARAGAT 507  
DB 554 AGGGAGGAATTTATGCGTTCGCTGCCAATTTGTTGTTATTTCTTTAGAGGGAATCAAGAT 613  
QY 508 GTGGTGCAGCTGGCAAGAGATCAAAATCTTCAATCAAGAGACGATCATTTTAAAGCTCC 567  
DB 614 ATAGTTGAATGGAGACGAATATGGAAGTTCAAAGTCAGATGGTCTTTCTTAAGCTCA 673  
QY 568 CTTGCATCTACTGCTTGTGTTCTATGACACACTCAAAACAAACAGATGCTCTCACTTCTC 627  
DB 674 CTTGCTTCTACTGCTGCTGTTTTTCATGCAACAGGAGAGCGGAAATGCGCTTGAATCTTG 733  
QY 628 AACTTCGTCTCAGCAAAATTTGGCGACTAGTTCCTTCCATTTACCCACTTGTATCTATTT 687  
DB 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGCTATCTCTGCGATCTGCTG 793  
QY 688 GAACGCTCTGGGCTGCTGCATACAGTTGAAACGCTTGGGAATCGATCGTATTTCAAGAAA 747  
DB 794 GAACGCTGTTGATCGTAGAATAATATGACGCTTGGAAATCTATAGACACTTTGAAAAG 853  
QY 748 GAAATCAAGAATCTCTGGATTAAGTTTATAGTACTTGGGACCGCCGAAAGAGCGGTGGGA 807  
DB 854 GAAATCAAGGAAGCTCTTGATTTATGATTATGTTTACAGGCAATTTGGA---CGAAGAGGAA 910  
QY 808 TGGGCAAGATGCAATCTTATCTCTGATGTCGATGACACTGCCCAATGGTCTTAGAATCTG 867  
DB 911 TGGGCGAGACTTAAATCCCATAGCAGATCTTGAACCACTGCTTTGGGATTTTCGATTTGCT 970  
QY 868 AGACTTTCATGATACAAATGTATCTTCAGATGTTTCTGAGAAATTTTCAGAGACGAGAAAGGA 927  
DB 971 CGGCTGATAGTGAATGTAATCTCCAGCAATTTTTCAGCAATTTCAAGATGCAATGGG 1030  
QY 928 GACTTCTTTTGTCTTTCGGTCAAACCGCAAAATTTGGTGTGACCGGATTAATCTTAACCTTAT 987  
DB 1031 AATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT 1090  
QY 988 AGATGTTCAAGTATGTTTTCGGGAGAAAAGATTAATGGAAGAGCTTAAGACCTTCACT 1047

Db 1091 AGAGCTTCCAGCTCGCATTTCCCGGAGAAACATTTCTTGATGAAGCTTAAAGCTTCGCT 1150  
Qy 1048 ACAATATCTCTCAAAATGCTTCTGCAAAACAGCAATTTGATGAAGTGGGCTGTCAAG 1107  
Db 1151 ACTAAATATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCAGTGCATGGGAACAACAA 1210  
Qy 1108 AGGATCTTCTCGAGAGGTGGATGCTATATAAGTATCCGTGGCATAGAAATATGCCA 1167  
Db 1211 CAAACCTGAGCCAGAGATCAATACGCGCTGAAGACTTCTTGGCATGCCAGTGTCCG 1270  
Qy 1168 AGATTGAGGCAAGAGTTTACATAGACCAATTTGGATCAAAATGATGTCTGGCTGGGAAG 1227  
Db 1271 AGAGTGAAGCAAGAGATAGTCTCAAGTGTATCGCCAGATTTATGACGCATAGCAAAA 1330  
Qy 1228 ACTGTGTATAGATGCTATATGTGACCAACGAAATATTTGGAGCTGGCCAAATTTGGAC 1287  
Db 1331 TGGGTTTACAGCTACCTACGTAAGCAATGAAGATTTTGGAGCTGGGAAATTTAGAT 1390  
Qy 1288 TTCAATATGCTGAGGCTTTACCAAAAGAGACTCAACACATTTGTACGCTGGTGAGA 1347  
Db 1391 TTCAACATTTACAGTCCATCCACCAAGAAAGAAATGAAGATGTTACCAAGCTGGTTT 1450  
Qy 1348 GAATCGGGATCAATGATCTTACATTTACCGCCAGCGGCTGTGGAAATGTTTCTCA 1407  
Db 1451 GATTCGGGGTTGCCACTATTTACCTTCGCTCGGAGAGCGGCTGGAAATTTCTACTTCT 1510  
Qy 1408 GTGCGGTTAGTATGTTTGGACCAAGATTCGCTCTTTGTAGAAATTTGCTATGCCAAGCT 1467  
Db 1511 GTAGCGGGGACCTATGAACCCAGTATGCCAAATGACAGTTTCCTCTTTACAAAAGTG 1570  
Qy 1468 TCTTCCCTCGCAGTTATTTAGACGATCTTTACGACCCACGAGTCTCTGGATGATCTT 1527  
Db 1571 GCATGCTTGCAGACTGTTCTGGACGATATGTATGACATTTATGACATAGTTTACAGAGTGGCTGGAG 1747  
Qy 1528 AAATGTTCTCTGAGCGGTCGAGAGTGGATATCTCTGTCTGGATAGGCTTCGGAT 1587  
Db 1631 AAGCTATTTACCTGAGGCTGTGAGAGATGGAGCTCTCTTTACAGAAACCTTCCAGAC 1690  
Qy 1588 AATCAGTTGAAAGTTTGTCTCTAGGGCTGTACAAACAGTGAATGGATTTGGAAAGAT 1647  
Db 1691 TAT---ATGAAACTATGTTTACCAATCTATTATGACATAGTTTACAGAGTGGCTGGAG 1747  
Qy 1648 GAATCAAGGAACAAGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGAGAGGCTTG 1707  
Db 1748 GCAGAGAGGAACAAGGGGCTGAAATTTGTCAGCTTTTTCAGAAAGGATGGAGATAT 1807  
Qy 1708 CTCGATCTGTATACCAAGAGCGAATGGTCCGAGCAAAAGTATGTGCGGACATTTCAAC 1767  
Db 1808 CTTCTGGGTTATATGAAGAGCTGAATGGTTAGCTGTGCTGATGTGCTTACCTTGGAC 1867  
Qy 1768 GAATATGTGGAAATGCCAAAGTGTCCATAGACATTTGGACAGTGTCTACTAACTCAATC 1827  
Db 1868 GAGTACATAAAGAAATGGAATCAATCTATCGGCAACAGTATATCTCTGTGATGGAGTG 1927  
Qy 1828 TTTTTCAC---TGGAGAAATTTACTTCTGATATATATTTTACAGCAAGTAGACCTTCGGTCC 1884  
Db 1928 TTGATATGAGTGGGCACTCTCTTTCCGAAGAGGCAATAGAGAAAGTAGATATCCAGGA 1987  
Qy 1885 AAATTTCTGC-----ATCTTGTGCTTTGACTGGAGCAATCAATCAATGACACCAAGACT 1938  
Db 1988 AGAGTGTCTCACAGAGCTGAATAGCCTCATTTCCCGCTGGCGGATGACAGAGACA 2047  
Qy 1939 TACCAGGCGGAGAAACCGTGTGTAATTTGTTTCCAGCGTACAGTGTCTACATGAGGAA 1998  
Db 2048 TATAAGCTGAGAGGCTCTGTGAGAAATTTGGCTCCAGCATTTGATGTTACATGAAAGAC 2107  
Qy 1999 AATCCGAGTGCACAGAGGAGAGCTCTAAGTCAATGTTTATGGTATCATCGAACGCA 2058  
Db 2108 CATCTGAATGTACAGAGGAGAGGCTCTCGATCACAATCTATAGCAATTTCTGGAGCGCG 2167  
Qy 2059 CTGAGGAATTTGAATTTGGAGTTGGCCAAACCCAGGAGCAATGCCCATTTGTGTGAGA 2118

Db 2168 GTGAGGAACCTGACAAAGAGAGTTTCTGAAGCCCG---ACGAGCTCCCATTTCCCTGCAAG 2224  
Qy 2119 AGACTGCTGTTCAACACTGCAGAGTGTATGAGCTGTGTTTATATATATGATCAGAGATGCTTTT 2178  
Db 2225 AAGATGCTTTTTCGAGGAGACAGAGTGGATGATATTTCAAGATGAGATGATTC 2284  
Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGCAGCCGAACCTCTTTTCGATCTGTGGC 2237  
Db 2285 GGTGTTTCCAAATTAGAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCC 2343

RESULT 8  
US-09-865-171-1  
; Sequence 1, Application US/09865171  
; Patent No. 6451576  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B  
; APPLICANT: Crock, John E  
; APPLICANT: Bohlmann, Jorg  
; APPLICANT: Steele, Christopher L  
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),  
; FILE REFERENCE: WSUR117468  
; CURRENT APPLICATION NUMBER: US/09/865,171  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/234,393  
; PRIOR FILING DATE: 1999-01-20  
; PRIOR APPLICATION NUMBER: 60/072,204  
; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Abies grandis  
US-09-865-171-1

Query Match 32.3%; Score 723.4; DB 4; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 3.8e-22;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;  
Qy 31 AATCCATCTGATATGATACAGCTTGGTGGCAAGAAATTCGTCATATGACGGCTCTGCT 90  
Db 134 ACTCATCTGCTTATGACAGCATGGTAGCGAGGGTGGCCGCAATTTGATGGCTCTGCT 193  
Qy 91 GCACCCCAATTTCCCAAAACGCTTCAATGGATTTCTGAACAATCACTGCCAGATGGCTCG 150  
Db 194 CGCCCGCAATTTCCCAAAACAGTTGACTGGATTTTGA AAAAACCAAGTTAAAGATGGTTCA 253  
Qy 151 TGGGTGAGGAGTGATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCCCTGCTC 210  
Db 254 TGGGGAATTTGAGTCCACTTTTCTGCTGTCGACCGCTTCTTGCCACTCTTTCTGTGTT 313  
Qy 211 CTCACCTCTCAAAATATGAATAAGGCGACATTTCAAGTCGAGAAAGGGGTTGAGTTTGTG 270  
Db 314 CTTGTGCTCTTAAATGGAAACGTTGGGATCTGCAAGTAGAGACGGAATTTGAATTCATA 373  
Qy 271 AGAAACACATCGAAGAAATGAAGCAAGCTGACAATCAAG---GCCAAGTGGATTC 327  
Db 374 AAGAGCAATCTGGAATCTAGTAAGGATGAAACCGATCAAGATAGCTTGTGTAAACAGACTT 433  
Qy 328 GAGGTGCTGTTTCTTCGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGATCTTCTCTAT 387  
Db 434 GAGATCATATTTCTTCTCTCTGTTTAAAGAAAGCTCAATCTCTGGCCTCGGACTTCCCTAC 493  
Qy 388 CACCTCCCTTTCTCTCCCAATCCACAAAGCGCAGAAAAGCTTCAAAAAGATTTCCC 447  
Db 494 GACCTCCCTTATATACATCTGTTGAGACTAAACGCGAGAAAGATTAGCAAACTTCA 553  
Qy 448 CTCAATGTTCTTTCATTAACCATCAGACGGGTTGCTCTACTCTCTGAGAGGTTTGCAGAT 507  
Db 554 AGGAGGAATTTATCGGTTTCGTCGCTCGCCATTTGTTGTTATTTCTTTAGAGGGAATACAGAT 613

```
QY 508 GTGGTGGACTGCGAAGAGTCAAAATCTTCAATCAAGAGCGGATCATTTTAAAGCTCC 567
Db 614 ATAGTGAATGGGAACGAATAATGGAAGTTCAAAGTCAGAGTGGGTCTTTCTTAAAGCTCA 573
QY 568 CTTGATCTACTGCTTGTGCTTCTATGACACTCAAAACAAACGATGCTCCACATTTCTC 627
Db 674 CTTGCTTCTACTGCTGCTGCTTCTATGACACAGAGAGCGGAATGCTTGATTTCTTG 733
QY 628 AACTTGGTCTCAGCAAAATTTGGCGACTACGTTCTTGGCATTTACCCACTTGATCTATTT 587
Db 734 AACAGTGTGATCAAGTTTGGAAATTTGTCTCCCTGCTCTATCTCTGATCTGCTG 793
QY 688 GAACCCCTCTGGCTCGATACAGTGAACCTTTGGGAATCGATCGCTATTTCAAGAA 747
Db 794 GAACCCCTCTGATCGTAGAATAATTTGACCCCTTGGGAATCTATAGACACTTTGAAAG 353
QY 748 GAAATCAAAAGATCTCTGATTAACGTTTATAGTACTGGGAGCGCGAAGAGCGGTGGGA 307
Db 854 GAATCAAGGAAGCTCTTGATTAATGTTTACAGGCATTTGAA---CGAAGAGGAATTTGG 310
QY 808 TGGGCAAGATGCAATCTTATCTGATGTCGATGACACTGCCATGGCTTTAGAAATCTTG 367
Db 911 TGGGCGAGACTAAATCCCATAGACAGATCTTGAGACCCTGCTTTGGGATTTGATTTGCTT 370
QY 868 AGACTTCATGATACATATGATCTTCAATGTTCTGGAGAAATTTCAAGACAGAGAAAGA 327
Db 971 CGGCTGCTAGTAGTCAATGATCTCCAGCCATTTTGTGACAACTTTCAAGAGTCCCAATGG 1030
QY 928 GACTTCTTTGCTTTGCGGTCAAAACGCAAAATTTGGTGTGACCGATAATCTTAACTTTAT 387
Db 1031 AAATTCATTTGCTCGACCGGTCAATTTCAAAAGATGTAGCAAGCATGCTGAATCTTTAT 1090
QY 988 AGATGTTCAAGATGATTTTTCGGGAGAAAAAGATAATGGAAGAGCTAAAGACCTTTCACT 1047
Db 1091 AGAGCTTCCCAGCTCGCATTTCCCGGAGAAAAACATTTCTTGATGAAGCTAAAGCTTGGCT 1150
QY 1048 ACAATATCTCCAAATGCTCTTGGCAAAAACAAGCATTTGATAAGTGGGCTGTCAAG 1107
Db 1151 ACTAAATATTTGAGAGAGAGCTTTGAGAAAAAGTGAGACTTCCAGTGCAATGGAACACAA 1210
QY 1108 AAGGATCTTCTGGAGAGGTGAGATGATCTATAAGATATCCGTGGCATAGAGTATGCCA 1167
Db 1211 CAAAACCTTGAGCAAGAGATCAATATACCGCTGGAAGACTTTTGGCATGCCAGTGTTCGG 1270
QY 1168 AGATGGAGGCAAGAGTTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAG 1227
Db 1271 AGAGTGGGAAGCAAGAGACTGTCAAGTGTATCGCCAGATTTATGACGCGATAGCAAA 1330
QY 1228 ACTGTTATAGATGCTATATGTAGCAACGAAAAATATTTGGAGCTGGCCAAAATTTGGAC 1287
Db 1331 TGGCTTTACAGCTACCTTACGTGAACCAATGAAAAAGTTTTAGAGCTGGGAAAAATTTAGAT 1390
QY 1288 TTCAATATGGTGCAGGCTTACACCAAGAGAGACTCAACATATGTCAAGCTGGTGGGA 1347
Db 1391 TTCAACATTTCCAGTCCATCCACCAAGAGAAATGAAGATGTTACCAAGCTGGTTTGA 1450
QY 1348 GAATCGGATTCATATGATCTTACATTTCACTCCCGCAGCGCTGTGGAATGATTTCTCA 1407
Db 1451 GATTCGGGTTGCCACTATTACCTTGGCTCGGAGAGCGCGCTGGAATTTACTTCTTA 1510
QY 1408 GTGGCGGTTAGTATGTTTGGAGCAGAAATTCGCTGCTGTGATGAATTTGCCATGCCAGACT 1467
Db 1511 GTAGCGGGGGGACCTATGAACCCAGTATGCAAAATGCAAGTCTCTCTTTACAAAAAGTG 1570
QY 1468 TCTTGGCTCGCAGTTATTTCTAGACCATTTTACGACACCCAGGATCTCTGATGATCTTT 1527
Db 1571 GCATGCTTGCAGACTGTTCTGACAGATATGTATGACACTTATGGAACCCCTAGATGAATG 1630
QY 1528 AAATGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGCTGATAGCGTTCGGGAT 1587
Db 1631 AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCTTTACAGAAAAACCTTCCAGAC 1690
QY 1588 AATCAGTTGAAAGTTTGTCTTCTAGGGCTGTGTAACAACACAGTGAATGGATTTGGAAAAAGAT 1647
```

```
Db 1691 TAT---ATGAAACTATGTTACCAAAATCTATTATGACATAGTTCCAGAGTGGCTTGGAG 1747
QY 1648 GCACTCAAGGAACAAGCCGCTGATGTGCTGGGTCTATCTTCGAAAAAGTATGGGAGGGCTTG 1707
Db 1748 GCAGAGAGGACACGGGCGTGAATTTGGTCAGCTTTTTCAGAAAGGATGGAGGATAT 1807
QY 1708 CTCGCATCTGTATACCAAGAGCCGAATGGTGGCGAGCAAGATGATGTGCCGACATTTCAAC 1767
Db 1808 CTTCTGGGTTATTTATGAAGAGCTGAATGGTTAGCTGTCTGAGTATGTGCTTACCTTGGAC 1867
QY 1768 GAATATGTGGAATAATGCCAAAGTGTCTCATAGCACTTTGGGACAGCTCTGCTACTAAATCAATC 1827
Db 1868 GAGTACATTAAGATGGAATCAATCTATCGGCCAAAGTATATCTTCTGTTGAGTGGAGTG 1927
QY 1828 TTTTTCAC---TGGAGAATTTACTTCTCTGATTAATTAATTTTACAGCAAGTAGACTTTGGGTCC 1884
Db 1928 TTGATAAATGGAATGGGCAACTCTCTTTTCGCAAGAGGCATTTAGAGAAAGTAGATTATCCAGGA 1987
QY 1885 AAATTTCTGC-----ATCTTGTGCTTTTGACTGGAAGCACTAATCAATGACACCAAGACT 1938
Db 1988 AGACGTGTTCTCACAGAGCTGAATAGCTCTCATTTCCCGCTGGCGGATGACACGAAAGACA 2047
QY 1939 TACCAGGCCGAGAGAAACCGTGTGTAATTTGGTTTCCAGCGTACAGTCTACATGAGGGAA 1998
Db 2048 TATAAAGCTGGAAGGCTCGTGGAGAAATTTGGCGTCCAGCAATTTGAATGTTTACATGAAAGAC 2107
QY 1999 AATCCGGAGTGCACAGAGAGAAAGCTCTAAAGTCAATGTTTATGTTATCATCGACCAACGCA 2058
Db 2108 CATCTGATGTACAGAGGAGAGGCTCTCGATCACTATAGCATTTCTGGAGCCGGCG 2167
QY 2059 CTGAAGGAATTCGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2118
Db 2168 GTGAAGGAAGTGCACAGAGAGTCTCTGAAGCCG---ACGACGCTCCCATTTGGCTGCAAG 2224
QY 2119 AGACTGCTTTCAACACTGCAAGAGTGTGAGTGTGTTTATATGTATGACAGAGATGGCTTT 2178
Db 2225 AAGATGCTTTTCGAGGAGACAGAGTGCAGTGGTGATATTTCAAGGATGGAGATGGATTC 2284
QY 2179 GGTATCTCTGACAAAGAGATGAAAGACCATGTTCAGCCGAACCTCTTTTCGATCTCTGTGGC 2237
Db 2285 GGTGTTTCCAAATTAGAGTCAAGATCATATCAAGAGAGTGTCTCATTTGAACCGCTGCC 2343
```

## RESULT 9

```
US-09-398-395A-45
; Sequence 45, Application US/09398395A
; Patent No. 6468772
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 64687721, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/398,395A
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(2347)
; OTHER INFORMATION: E-alpha-bisabolene synthase
```



US-09-398-395A-45									
Query Match 32.3%; Score 723.4; DB 4; Length 2424; Best Local Similarity 59.8%; Pred. No. 3.8e-222; Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;									
QY	31	AATCCATCTGCATATGATACAGCTTGGTGGCAGAAATTCGGTCAATTTGACGGCTCTGGT	50	QY	1048	ACAAATCATCTCCAAATGCTCTTTGCCAAAAACAACGCAITTTGATAAGTGGGCTGTCAAG	1107		
DB	134	ACTCCATCTGCTTATGACACAGATGGGTAGCAGGGTGCCGCCATTTGATGGCTCTGCT	193	DB	1151	ACTAAATATTTGAGAGAGCTCTTTGAGAAAGTGCAGACTTCCAGTGCATGGAACAACAAA	1210		
QY	91	GCACCCAAATTTCCCAACGGCTTCAATGGAATCTGAACAATCAACTGCCAGATGGCTCG	150	QY	1108	AAGGATCTTCTCTGGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAAGTATGCCA	1167		
DB	194	CGCCCGCAATTTCCCAACAGATTGACTGGATTTGAAACACCAAGTTAAAGAGATGGTTCA	253	DB	1211	CAAAACCTGAGCCAAAGAGATCAAAATACGCGCTGAAGACTTCTTTGGCATGCCAGTGTTCG	1270		
QY	151	TGGGGTGAGGAGTGCATTTTCTGGCGGTATGACAGAGTATTTAAACACTCTCGCTGCCTC	210	QY	1168	AGATTGGAGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAG	1227		
DB	254	TGGGGAATTCAGTCCCACTTCTGCTGTCGACCGCTCTTCTTGGCACTCTTCTTGTGTT	313	DB	1271	AGAGTGGAGCAAGAGATACCTGTCAAGTGTATGCCCGAGATTTATGCAACATAGCAAAA	1330		
QY	211	CTCACTCTCAAAATATGGAATAGGGCGACATTCAGTGCAGAAAGGGTTGAGTTGTG	270	QY	1228	ACTGTGTATAGATGCTATATGTGAGCAACGAAAAATATTTGAGCTGGCCAAATTTGGAC	1287		
DB	314	CTTGTGCTCTTAATAGGAAGCTTGGGGATCTGCAAGTAGAGCAGGAAATGAATTCATA	373	DB	1331	TGCGTTTAAACAGCTACCCCTACGTGAAACAAATGAAAGTTTTTTAGAGCTGGGAAAAATTAG	1390		
QY	271	AGAAAAACATGGAAGAAATGAAGAGCAAGCTGACAAATCACAG---GCCAAGTGGATTC	327	QY	1288	TTCAATATGGTGAGGCTTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGGTGAGA	1347		
DB	374	AAGAGCAATCTGGAACTAGTAAGAGATGAAACCGATCAAGATAGCTTTGGTAACAGACTTT	433	DB	1391	TTCAACATTTTCCAGTCCATCCACCAAGAAAGAAATGAAGATTTTACCAGCTGGTTTAGA	1450		
QY	328	GAGGTGCTGTTTCTGCAATTTAGATGAAGCAAAAGCTTTGGATTTGGATCTTCTTTAT	387	QY	1348	GAATCGGGATTTCAATGATCTTACATTCACCCGCGAGCGGCTGTGGAAATGTATTTCTCA	1407		
DB	434	GAGATCATATTTCTTCTCTGTTAAGAGAAGCTCAATCTCTGCGCCTCGGACTTCCCTAC	493	DB	1451	GATTCGGGGTTGCCACTATTTCACCTTCGCTCGGAGAGGCGCTGGAAATTTCTACTTCTTA	1510		
QY	388	CACCTCCCTTTTCATCTCCCAATCCACAAAGCGCCAGAAAAAGCTTCAAAAGATTTCC	447	QY	1408	GTGCGGTTAGTATGTTTGGCCAGAAATTCGCTGCTGTGAGAAATGCTTATGCCAAGACT	1467		
DB	494	GACCTGCTTATATACATCTGTCGAGACTAAACGGCAGGAGAAATTAGCAAACTTTTCA	553	DB	1511	GTAGCGCGGGGACCTATGAACCCAGTATGCCAAATGCAGGTTCTCTTTTACAAAAGTG	1570		
QY	448	CTCAATGTTCTTCAATCCATCAGACGCGTGTCTACTCTCTGAGGGTTTGGCAAGAT	507	QY	1468	TCCTTGGCTCGCAGTTATTTCTAGACGATCTTTTACGACACCCACGGATCTCTGGATGATCTT	1527		
DB	554	AGGAGGAAAAATTTATGCGGTTCCGTCGCAATTTGTATTTCTTTAGAGGGAATACAAGAT	613	DB	1571	GCATGCTTGCAGACTGTTCTGAGCGATATGTATGACACTTTATGGAACCTTAGATGAAATG	1630		
QY	508	GTGTTGAGTGGCAGAGATCAAAATCTTCAATCAAGAGACGGATCAATTTTAAAGCTCC	567	QY	1528	AAATTTGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGTTCCGGAT	1587		
DB	614	ATAGTTGAATGGGAACGAATATGGAAGTTTCAAGTCAAGATGGGTCTTTCTTAGCTCA	673	DB	1631	AAGCTATTTACTGAGGCTGTGAGAAAGTGGGACCTCTCTTTTACAGAAAACCTTCCAGAC	1690		
QY	568	CTGCTACTCTGCTGTGCTTCTCATGCACACTCAAAAACAACGATGCTCCACTTTCTC	627	QY	1588	AATCAGTTTGAAGTTTTCCTTAGGGCTGTACAACACAGTGAATGGATTTTGGAAAAAGAT	1647		
DB	674	CTGCTTCTACTGCTGCGTTTTCATGCACAGGAGACGGAAATGCTTGAATTTCTTG	733	DB	1691	TAT---ATGAAACTATGTTTACCAATCTTATATGACATAGTTTACGAGGTGGCTTGGAG	1747		
QY	628	AACTTGCTGCTCAGAAATTTTGGCGACTACGTTCTCTGCCATTTACCACTTGAATCTATT	687	QY	1648	GGACTCAAGGAACAGGCGGTGATGTGCTGGGCTATCTTTCGAAAAGTATGGGAGGCTTG	1707		
DB	734	AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGTATCTCTGTGGATCTGCTG	793	DB	1748	GCAGAGAAGGAACAGGCGGCGTGAATTTGGTCACTGCTTTTTCAGAAAAGGATGGAGGATTTAT	1807		
QY	688	GAACGCTCTGGGCTGTGATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAAA	747	QY	1708	CTCGCATCTGTATACAAAAGACCGAATGGTGGCAGCAAAAGTATGTGCCGACATTTCAAC	1767		
DB	794	GAACGCTCTTGATCGTATAGATTAATTTGACGCTTGGAACTATAGACACTTTGAAAG	853	DB	1808	CTTCTGGGTTTATATGAAGAGCTGAATGGTTAGCTGTGCTGAGTATGTGCTTACCTTGGAC	1867		
QY	748	GAATCAAGAAGATCTCTGATTTACGTTTATAGGTACTGGAGCCGCAAGAGCGCTGGGA	807	QY	1768	GAATATGTGGAATATGCCAAAGTGTCCATAGCACCTTGGGACAGTCTGTACTTAACTCAATC	1827		
DB	854	GAATCAAGAAGCTCTTGATTTATGTTTACAGGCAATGGAA---CGAAAAGAGAAATGGG	910	DB	1868	GAGTACATAAAGAATGGAATCACTATCTATCGGCCAACGTATATCTTCTGTGTAGTGGAGTG	1927		
QY	808	TGGGCAAGATGCATCTTATCTCTGATGTGATGACACTGCCATCGGCTTTAGAAATCTCTG	867	QY	1828	TTTTTTCAC---TGGAGAAATTTCTCTGATTTACATTTTACAGCAAGTACAGCTTCCGGTCC	1884		
DB	911	TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTGATTTGCTT	970	DB	1928	TTGATAATGGATGGCGCAACTCTTTTCGCAAGAGGCATTTAGAGAAAGTAGATTTATCCAGGA	1987		
QY	868	AGACTTTCATGGATACAATGATATCTTTCAGATGTTCTTGGAGAAATTTTCAGAGACGAGAAAGGA	927	QY	1895	AAATTTCTGCG-----ATCTTGTGCTTTTGAATGGAACGACTAATCAATGACACCAAGACT	1938		
DB	971	CGGCTGATAGGATCAATGATATCTCCAGCCATTTTTCAGCAACTTTCAAGATGCCAATGGG	1030	DB	1988	AGACGCTTCTCACAGAGCTGAATAGCCTCATTTTCCGCTGGCGGATGACACGAGAGACA	2047		
QY	928	GACTTCTTTTGTCTTGGCGGTCAACGCAAAATGGTGTGACCGATTAATCTTAACTTTTAT	987	QY	1939	TACCAAGCGGAGAGAAACCGTGTGATTTTTCAGCGGTACAGTGTCTACATGAGGGAA	1998		
DB	1031	AAATTTCAITTTGCTCGACCGGTCAATTTCAACAAGATGTAGCAAGCATGTCTGAATCTTTAT	1090	DB	2048	TATAAGCTGGAAGGCTCGTGAGAAATTTGGCGTCCAGCAATTTGAATGTTTACATGAAGAC	2107		
QY	988	AGATGTTCAAGATATGTTTTCGGGAGAAAGATAATGGAAGAGCTTAAGACCTTCACT	1047	QY	1999	AATCCGGAGTGCACAGAGGAAGAGCTCTTAAGTCAATGTTTATGGTATCATCTCGACAAGCA	2058		
DB	1091	AGAGCTTCCAGCTGCGATTTTCCGGAGAAACAATTTCTTGTATGAAGCTTAAAGCTTCGCT	1150	DB	2108	CATCTGATGTACAGAGGAGAGGCTCTCGATCACATCTATAGCATTTCTGGAGCCGCG	2167		







Db 1031 AAATTCATTGCTGCACCGGTCAATTCAACAAGATGTAGCAAGCATGCTGAATCTTTAT 1090  
Qy 988 AGATGTTCAAGTATGTTTCCGGAGAGAAAGATAATAGGAAGAGCTTAAGACCTTCACT 1047  
Db 1091 AGAGCTTCCAGCTCGCATTTCCCGAGAGAAACATTTCTGATGAAGCTAAAGCTTCGCT 1150  
Qy 1048 ACAATCATCTCCAAATGCTTTCGCCAAACACGCAATTTGATAAGTGGGCTGTCAAG 1107  
Db 1151 ACTAAATATTTGAGAGAGCTCTTGAGAAAGTGAGACTTCCAGTGCAATGGAAACACAA 1210  
Qy 1108 AAGGATCTTCTCGAGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAATGCA 1167  
Db 1211 CAAACCTGAGCCAGAGATCAATACGCGCTGAAGACTTCTTGGCATGCCAGTGTTCG 1270  
Qy 1168 AGATTGAGGCAAGATTTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGAG 1227  
Db 1271 AGAGTGAAGCAAGAGATACTGTCAAGTGTATCGCCAGATATGACGATAGCAAA 1330  
Qy 1228 ACTGTATAGATCTATATGTAGCAACGAAATATTTGGAGCTGCCAAATGGAC 1287  
Db 1331 TGCCTTTACAGCTACCTCACTGCTGAGCAATGAAGATTTTATAGAGCTGGGAAATTAGAT 1390  
Qy 1288 TTCAATATGTTGCGAGCTTACACCAAAAGGAGACTCAACATTTGTCAAGCTGGTGGA 1347  
Db 1391 TTCAATATTCAGTCACTCAACCCAGAGAAATGAAGATGTTACAGCTGGTTTGA 1450  
Qy 1348 GAATGGGATTCATGATCTTACATTCACCCCGCAGCGCTGTGGAAATGATTTCTCA 1407  
Db 1451 GATTCGGGTTGCCACTATTACCTTCGCTCGGAGAGCGCTGGAATTTCTACTTTA 1510  
Qy 1408 GTGGGGTTAGTATGTTGAGCCAGAAATTCGCTGCTGTAGAAATGGCTATGCCAGACT 1467  
Db 1511 GTAGCGCGGGGACCTATGAACCCAGTATGCCAAATGCAAGTTCCTCTTTACAAAAGTG 1570  
Qy 1468 TCTTGCCTCGCAGTTATTCTAGACCATCTTTACGACACCCAGGATCTCTGATGATCTT 1527  
Db 1571 GCATGCTTGACAGTGTCTGACACATATGTATGACACTTATGGAACCTAGTGAATG 1630  
Qy 1528 AAATGTTTCTCGAAGCGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGCTTCGGAT 1587  
Db 1631 AAGCTATTCTAGGCTGTGAGAGATGGGACCTCTCTTTACAGAAACCTTCCAGAC 1690  
Qy 1588 AATCAGTTGAAATGTTCTCTAGGCTGTACAAACAGATGAATGGATTTGGAAAGAT 1647  
Db 1691 TAT---ATGAAACTATGTTACCAATCTATTATGACATAGTTTACGAGGTGCTTGGAG 1747  
Qy 1648 GGACTCAAGGAACAGGCGCTGATGCTGTGGCTATCTTCGAAAGATAGGAGGCTTG 1707  
Db 1748 GCAGAAAGGAACAGGCGGTGAATTTGTGAGCTTTTTCAGAAAGGATGGGAGTAT 1807  
Qy 1708 CTCGCATCGTATACCAAGAGCCGAATGGTGGCAGCAAAAGTATGTGCCGACATTTCAAC 1767  
Db 1808 CTTCTGGGTTATATGAAGAGCTGAATGGTTAGCTGTGATGATGTGCCTACCTTGGAC 1867  
Qy 1768 GAATATGGAAGAAATGCCAAAGTGTCCATAGACATTTGGCGACAGTGTACTAACTCAATC 1827  
Db 1868 GAGTACATAAAGATGAATGACATCTATCGCCCAACGCTATATCTCTGTGAGTGGAGTG 1927  
Qy 1828 TTTTTCAC---TGGAGAAATTAATCTGATTTACATTTTACAGCAAGTAGACCTTCGTTCC 1884  
Db 1928 TTGATAATGGATGGGCAACTCTCTTCGCAAGAGGCAATTAGAAGATGATTTACAGGA 1987  
Qy 1885 AAATTTCTGC-----ATCTGTGTCTTTGACTGGACGACTAATCAATGACACCAAGACT 1938  
Db 1988 AGACGTGTTCTCAGAGAGCTGAATAGCTCATTTCCCGCTGGCGGATGACACGAGACA 2047  
Qy 1939 TACAGGCGCGAGAGAAACGTTGGTGAATTTGGTTTCCAGGTTACAGTGTACTATGAGGAA 1998  
Db 2048 TATAAAGCTGAGAGGCTCGTGAGAAATGGCGTCCAGCATTTGAATGTTTACATGAAGAC 2107  
Qy 1999 AATCGGAGTGCACAGAGAGAGAGCTCTAGTCAATGTTATGTTATCATCATGCAACGCA 2058  
Db 2108 CATCTGAATGTACAGAGGAAGAGGCTCTCGATCATCTATAGCAATTTCTGGAGCGGCG 2167

Qy 2059 CTGAAGAAATGAATTTGGAGTTGGCCAAACCCAGCGAGCAATGCCCAATTTGTGTGAGA 2118  
Db 2168 GTGAAGAAACTGACAGAGAGTTTCTGAAGCCCG---ACGACGTCCCAATTCGCTGCAAG 2224  
Qy 2119 AGACTGCTCTTCAACACTGCAAGAGTGATGCAGCTGTTTATATGTACAGAGATGGCTTT 2178  
Db 2225 AGATGCTTTTCGAGGAGCAAGAGTGACGATGGTATATTCGAAGATGGAGATGGATTC 2284  
Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGAGCCGAACCTCTTTTCGATCTCTGTGGC 2237  
Db 2285 GGTGTTTCCAAATTAGAAGTCAAGATCATATCAAGAGGTGTCTCATTTGAACCGCTGCC 2343

## RESULT 12

US-09-903-012B-45  
; Sequence 45, Application US/09903012B  
; Patent No. 6569656  
; GENERAL INFORMATION:  
; APPLICANT: Chappel, Joseph  
; APPLICANT: No. 65696561, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/903,012B  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(2347)  
; OTHER INFORMATION: E-alpha-bisabolene synthase  
US-09-903-012B-45

Query Match 32.3%; Score 723.4; DB 4; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 3.8e-222;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

Qy 31 AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCGTCATTTGACGGCTCTGGT 90  
Db 134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGAGGTGCCGCCATTGATGGCTCTGCT 193  
Qy 91 GCACCCCAATTTCCCAACCGCTTCAATGATTTGAACTCAATCACTGCGCAGATGGCTG 150  
Db 194 CCGCCGCAATTTCCCAACAGTTGATGATTTGAAAACAGTTAAAGATGGTTCA 253  
Qy 151 TGGGGTGGAGGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCTC 210  
Db 254 TGGGGAATTCAGTCCCACTTTCTGCTGCGACCGTCTTCTGCCACTCTTTCTTGTT 313  
Qy 211 CTCACCTCTCAAAATATGGAATAAGGCGGACATTCAGAGTCAGAAAAGGGTTGAGTTGTG 270  
Db 314 CTTGTGCTCTTAAATGGAACGTTGGGATCTGCAAGTAGAGCAGGAAATGAATTCATA 373  
Qy 271 AGAAAACATCGAGAGAAATGAAGGACGAGCTGACAACTACAG---GCCAGTGGATTC 327  
Db 374 AAGAGCAATCTGGAACCTAGTAAAGGATGAAACCGATCAAGATAGCTTTGGTAAACAGACTTT 433  
Qy 328 GAGGTGCTGTTTCTCGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATCTTCTTAT 387  
Db 434 GAGATCATATTTCTTCTCTCTGTTAAGAGAGCTCAATCTCTGCGCTTCGAGCTTCCCTAC 493

QY 388 CACCTCCCTTTCATCTCCCAATCCCAAAAGCGCCAGAAAAGCTTCAAAAGATTCCC 447  
Db 494 GACCTGCCCTTATATACATCTGTTGCGAGCTAAACCGCAGGAAAGATTAGCAAACTTTCA 553  
QY 448 CTCAAATGTTCTTCAATCAACATCAGACGGCTGTCTTACTCTCTGAGGGGTGTGCAAGAT 507  
Db 554 AGGAGGAATTTATGCGGTTCGTCGCCATTTGTTATCTTTAGAGGGAATACAAAGAT 613  
QY 508 GTGGTGGAGTGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTTAAGTCC 567  
Db 614 ATAGTTGAAATGGGAACGAATTAATGGAAGTTCAAAAGTCAGGATGGGTCTTTCTTAAAGCTCA 673  
QY 568 CTTGCACTACTCTGTTGCTTCAATGACACATCAAAACCAACGATCCCTCACTTCTC 627  
Db 674 CTTGCTTCTACTGCTGCTGTTTTCATGCAACAGGAGACGGGAATGCCCTTGAATCTTGT 733  
QY 628 AACTTCTGCTGCTAGCAAAATTTGGCGACTAGCTTCTTTGCCATTAACCACTTGATCTATTT 687  
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTGTTCCCTGCTGTATCTGTGATCTGCTG 793  
QY 688 GAACGCTCTGGGCTGTGATGATCAAGTTTGAAGCTTGGGAATCGATCGTATTTCAAGAA 747  
Db 794 GAACGCTGTTGATCGTAGATAATTTGATGACGCTTGGAACTTATAGACACTTTGAAAG 853  
QY 748 GAAATCAAGNATCTCTGCTTACGTTTATAGTACTGGGACGCGGAAAGGCGTGGGA 807  
Db 854 GAAATCAAGGAGCTCTTGTATTTGATTTTACAGGCAATTTGAA -- CGAAAGAGGAATTTGGG 910  
QY 808 TGGGCAAGATGCAATCTTATCTCTGATGTCGATGACACTGCCATGGGTCTTGAATCTGT 867  
Db 911 TGGGCAAGACTAATCCCATAGCAGATCTTGGACCACTGCTTTGGGATTTTGGATGCTT 970  
QY 868 AGACTCATGATCAATGATCTTTCAGATGTTCTCGGAAATTTTCAGAGACGAGAAAGGA 927  
Db 971 CGGCTGATAGTAAATGATCTCTCCAGCACTTTTGAACACTTTCAAGAGATGCCAATGGG 1030  
QY 928 GACTTCTTTTGGCTTGGCGTCAAAAGCAATTTGGTGTGACGATTAATCTTAACTTTAT 987  
Db 1031 AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCTGCTGAACTTTAT 1090  
QY 988 AGATGTTTCAAGATGATTTTCCGGGAGAAAAGATAATGGAAGAGCTTAAGACCTTCACT 1047  
Db 1091 AGACTTCCAGCTCGCATTTCCCGGAGAAAACATTTCTGATGAAGCTTAAAGCTTCTGCT 1150  
QY 1048 ACAATCATCTCAAAATGCTTTCGCAAAACACAGCATTTGATGAGTGGCTGTCAAG 1107  
Db 1151 ACTAAATATTTGAGAGAGCTCTTTGAGAAAAGTGAGACTTCCAGTGCATGGAACAA 1210  
QY 1108 AAGATCTTCTGAGAGGTGGAGTATGCTATAAGTATCCGTGGCATAGAGTATGCCA 1167  
Db 1211 CAAACTGAGCCAAAGATCAATACGGCTGAGACTTCTTGGCATGCCAGTGTTCGG 1270  
QY 1168 AGATTGGAGGCAAGATTATAGAGCAATTTGGATCAATGATGTCTGGCTGGGGAAG 1227  
Db 1271 AGAGTGAAGCAAGAGATACCTGTCAAGTGTATCGCCAGATTTATGCAAGCATAGCAAA 1330  
QY 1228 ACTGTGATAGATCTATATGTGAGCAAGCAAAATATTTGAGCTGGCCAAATTTGGAC 1287  
Db 1331 TGCGTTTACAGCTACCTTACCTGTGAGCAATGAAAAGTTTTTATAGAGCTGGGAAATTAGAT 1390  
QY 1288 TTCATATGGTGCAGGCTTACCAAAAGGAGACTCAACATTTGTCAAGCTGGTGGAGA 1347  
Db 1391 TTCACATTTCCAGTCCATCCACCAAGAGAAATGAAATGTTTACCAGCTGGTTTGA 1450  
QY 1348 GAATCGGATCAATGATCTTTCATTTACCCCGCAGCGCTGTGGAATGATTTCTCA 1407  
Db 1451 GATTGGGGTTCGCACTATTTCACTTCTGCTCGGAGAGCGCTGGAATTTCTACTTTTA 1510  
QY 1408 GTGGGCTTGTATGTTTGGAGCAGAAATTTGCTGCTTGTAGAAATTCCTATGCCAAGACT 1467  
Db 1511 GTAGCGGGGAGCTATGAAACCCCAATGCAATGCAAGTCTCTCTTTTACAAAGTG 1570  
QY 1468 TCTTCCCTCGCAGTTATTTTAGACGATCTTTAGCAACCCAGGATCTCTGAGATGATCTT 1527

Db 1571 GCATGCTTGACAGACTGTTCTGGACGATATGTATGACACTTATGGAACCTTAGATGAATTTG 1630  
QY 1528 AAATGTTCTTGAAGCGTCCGAGATGGGATATCTCTGTCTGATAGGCTTCCGGAT 1587  
Db 1631 AAGCTATTCATGAGGCTGTGAGAGATGGGACCTCTCTCTTTACAGAAACCTTCCAGAC 1690  
QY 1588 AATCAGTTGAAAGTTTCTTCTAGGCTGTACAAACAGAGTAAATGGAATTTGGAAAAGAT 1647  
Db 1691 TAT --- ATGAACTATGTTACCAATCTATTTATGACATAGTTTACGAGGTGCTTGGGAG 1747  
QY 1648 GGACTCAAGGAACAGGCGGTGATGTCTGGGCTATCTTCAAAAGATATGGGAGGCTTTG 1707  
Db 1748 GCAGAGAAGGAACAGAGGCGGTGAATTTGGTCACTTTTTCAGAAAGGATGGGAGGATTA 1807  
QY 1708 CTCGCATGATATACCAAGAGCGAATGTTGGCGAGCAAAAGTATGTGCCGACATTTCAAC 1767  
Db 1808 CTTCTGGGTTATTTAGAAAGCTGAATGGTTAGCTGTGATGATGTGCTTACCTTTGGAC 1867  
QY 1768 GAATATGTGAAAATGCCAAAAGTGTCCATAGCACTTGGGACAGTCTGATAAATCAATC 1827  
Db 1868 GAGTACATAAAGATGGAATCACATCTATCGCCAAACGTATATCTTCTGTGAGTGGAGTG 1927  
QY 1828 TTTTTCAC --- TGAGAAATTTACTTCTGATTTACATTTTACAGCAAGTAGACCTTCGGTCC 1884  
Db 1928 TTGATAATGGATGGGCAACTCTCTTTTCCGAAGAGGCAATTAGAGAAAGTAGATTTATCCAGGA 1987  
QY 1885 AAATTTCTGC --- ATCTTGTGCTTTGACTGGACGACTAATCAATGACACCAAGACT 1938  
Db 1988 AGAGTGTCTTCAACAGAGCTGATAGCTCTATTTCCGCTGGCGATGACAGGAGACA 2047  
QY 1939 TACCAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGGTCACAGTGTACATGAGGGAA 1998  
Db 2048 TATAAGCTGAGAGAGGCTCGTGGAGAAATTTGGCTCCAGCAATTTGAATTTTACATGAAAGAC 2107  
QY 1999 AATCCGGAGTGACAGAGAGAGAGCTCTAGTCTGATTTTATGTTATGATCATCGACACGCA 2058  
Db 2108 CATCTGAAATGTACAGAGAGAGAGCTCTCGATCATCATCTATAGCATTTCTGGAGCCGGC 2167  
QY 2059 CTGAAGGAATTAATTTGGAGTTTGGCCAAACCCAGGAGCAATGCCCCATTTGTGTGAGA 2118  
Db 2168 GTGAAGGAATGTACAGAGAGAGTTTCTGAAGCCG --- ACAGCTGCCATTTCCGCTGCAAG 2224  
QY 2119 AGACTGCTGTTCAACACTGCAAGAGATGATGAGCTGTTTTTATATGTATGACAGAGATGCTTT 2178  
Db 2225 AAGATGCTTTTCCAGGAGACAGAGAGTGCATGTTGATATTTCAAGGATGGAGATGATTC 2284  
QY 2179 GGTATCTCTGCAAAAGAGATGAAGACCATGTGACCGAACTCTTTTTCGATCTCTGTGGC 2237  
Db 2285 GGTGTTTCCAAATTTAGAGTCAAGATCATATCAAAAGAGTGTCTCATTTGNAACCGCTGCC 2343

RESULT 13  
US-09-900-797-45  
; Sequence 45, Application US/09900797  
; Patent No. 6645762  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 66457621, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0

```

; SEQ ID NO 45
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(2347)
; OTHER INFORMATION: E-alpha-bisabolene synthase
US-09-900-797-45

Query Match      32.3%; Score 723.4; DB 4; Length 2424;
Best Local Similarity 59.8%; Pred. No. 3.8e-222;
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

QY 31 AATCCATCTGCATATGATCAGCTTGGGTGGGCAAGAAATTCGGTCAATTAAGCGCTCTGGT 90
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGGGTGGCCGCCATTCATGGCTCTGCT 193
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 91 GCACCCCAATTTCCCAACCGCTTCAATGATTTCTGAACAATCAACTGCGCAGATGGCTCG 150
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 194 CGCCCGCAATTTCCCAACAGTTGATGATTTTGAAGAACCCAGTTAAAGATGGTTCA 253
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 151 TGGGTGAGGAGTGCAATTTTCTGGCGTATGACAGAGTTTAAACAATCTCTCGCTCGCTC 210
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 254 TGGGGAATTCAGTCCCACTTCTGCTGTCGACCGTCTTCTGCCACTCTTCTTGTGTT 313
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 211 CTCACCTCTCAAAATATGGAATTAAGGCGGACATTCAGAGTCAGAAAGGGGTGATTTG 270
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 314 CTGTGCTCTTAAATGGAACGTTGGGATCTGCAAGTAGAGCGGGAATTTGAATTCATA 373
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 271 AGAAACACATCGAGAAATGAAGGACGAGCTGACAAACACAG---GCCAGTGGATTC 327
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 374 AAGAGCAATCTGGAATCTAAGAGGATGAACCGATCAAGATAGCTTGGTAAACAGACTTT 433
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 328 GAGGTGCTGTTTCTGCAATGTTAGATGAAGCAAAAGCTTGGGATTTGGATCTTCTTAT 387
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 434 GAGATCAATTTCTTCTCTGTTAAGAGAGCTCAATCTCTCGCGCTCGGACTTCCCTAC 493
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 388 CACCTCCCTTTCACTCTCCAAATCCACAAAGCGCAGAAAGCTTCAAAAGATTTCC 447
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 494 GACCTGCTTATATACATCTGTTCAGACTAAACCGCAGGAAAGATTAGCAAACTTTCA 553
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 448 CTCATGTTCTTCATTAACCATCAGAGCGGTGCTCTACTCTCTGAGGGTTGCAAGAT 507
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 554 AGGGAGGAATTTATGCGGTTCGTCGCCATTTGTTATTTCTTTAGAGGAATACAAAGAT 513
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 508 GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGAGCGGATCAATTTTAAGCTCC 567
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 614 ATAGTTGAATGGGACGATATATGGAATTCAGAGTCAGATGGGTCTTCTTTAAGCTCA 573
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 568 CTGCACTACTGCTTGTCTTTCATGACACTCAAAACAAACAGATGCTCCACTTTCTC 527
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 674 CTTGCTTCTACTGCTGCTTTTCATGACACAGAGAGCGGAAATGCTTGAATTTCTTG 733
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 628 AACTTCGCTCAGCAAAATTCGGGACTAGCTTCTTCTGCAATTCACCACTTGAATTT 587
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 734 AACAGTGATGATCAAGTTTGAATTTTGTTCCTGCTGCTGATCTCTGCTGCTGCTG 793
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 688 GAAACGCTCTGGCTGTCGATACAGTTGAACGCTTGGGAATCGATCGCTATTTTCAAGAA 747
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 794 GAAACGCTGTTGATCGTAGATATATTTGACGCTTGGGAATCTTAGACACTTTGAAAG 353
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 748 GAAATCAAGAAATCTCTGATTAAGTTTATAGTACTGGGACCGCGAAAGAGCGGTGGGA 307
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 854 GAAATCAAGAAAGCTCTTGATTTATGTTTACAGGCAATTTGAA---CGAAAGGAATTTGG 310
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 808 TGGGCAAGATGCAATCTTATCTGATGTCGATGACACTGCGCATGGCTTTTGAATCTCTG 367
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 911 TGGGCGAGACTAAATCCCAATAGCAGATCTTGAGACCACTGCTTTGGGATTTTCAATGCTT 370
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 868 AGACTTCTAGATACATGATCTTCTGAGATGTTCTGAGAAATTTTCAAGACGAGAAAGGA 327
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 971 CGGCTGCATAGGTACAAATGTAATCTCCAGCAATTTTTCACAACTTTCAAGATGCCATGGG 1030
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

QY 928 GACTTCTTTTGTCTTTCGGGTCAAAACGCAAAATTTGGTGTGACCGGATTAATCTTAACTTTAT 987
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1031 AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT 1090
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 988 AGATGTTCAACAGTATGTTTTCGGGAGAAAGATTAATGGAAGAGTAAAGCATTTCACT 1047
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1091 AGAGCTTCCAGCTCGCATTTCCCGGAGAAAACATTTCTTGATGAAGCTTAAAGCTTCGCT 1150
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1048 ACAATATCATCTCCAAAATGCTTTGCAAAAACAAACGCAATTTGATAGTGGCTGTCAAG 1107
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1151 ACTAATATTTGAGAGAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAACAACAAA 1210
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1108 AAGGATCTTCTCGGAGAGGTGGAGTATGCTATTAAGATATCCGTGGGCATAGAAGTATGCCA 1167
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1211 CAAAACCTTGAGCCCAAGAGATCAAAATAGCGCTCAAGAACTTCTTGGCATGCGAGTGTCCG 1270
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1168 AGATTGAGGCAAGAAAGTTACATAGAGCAATTTTGGATCAATATGATGCTCGCTGGGAG 1227
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1271 AGAGTGAAGCAAGAGATATCTGTCAAGTGTATCGCCAGATTTATGCACGCATAGCAAAA 1330
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1228 ACTGTGTATAAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGCCAAATTCGAC 1287
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1331 TCGGTTTACAGCTACCTTACGTGAACATGAAGTTTTTTAGAGCTGGGAAATTTAGAT 1390
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1288 TTCAATATGGTGCAGGCTTTACACAAAAGAGAGCTCAACACATTTGTTCAGCTGGTGAGA 1347
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1391 TTCAACATTTATCCAGTCCATCCACCAAGAAAGAAATGAAGAAATGTTACAGCTGGTTAGA 1450
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1348 GAATCGGGAATTCATGATCTTACATTCACCGCAGCGGCTGTGGAATTTGTTCTCA 1407
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1451 GATTCGGGTTGCGCACTTATTCACCTTCGCTCGGAGAGCGCTGGAAATTTCTACTCTTA 1510
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1408 GTGGCGTTAGTATGTTTGGAGCAGAAATTCGCTGCTTGTAGAATTTGCTATGCAAGACT 1467
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1511 GTAGCGCGGGGACCTATGAACCCAGTATGCCAAATGCAGGTTCTCTTTTACAAAAGTG 1570
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1468 TCTTGCTCGCAGTATTTCTTACAGCATTTTACGACACCCACCGATCTCTGGATGATCTT 1527
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1571 GCATGCTTCAGACTGTTCTGAGCATATGATGACACTTATGGAACCTTAGATGAATG 1630
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1528 AAATGTTCTCTGAAGCGGTCCGAGATGGAATCTCTGTGCTGGATAGCTTCCGGAT 1587
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1631 AAGCTATTCACTGAGGCTGTGAGAAATGGGACCTCTCTTTTACAGAAAACCTTCCAGAC 1690
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1588 AATCAGTTCAAAAGTTTGTCTTCTAGGGCTGTACAAACAGTGAATGGAATTTGGAAGAAT 1647
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1691 TAT--ATGAAACTATGTTACCAATCTATATGACATAGTTCAAGAGGTGGCTTGGAG 1747
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1648 GGACTCAAGGAACAAGCGGTGATGTGGGCTATCTTCGAAAAAGTATGGGAGGGCTTG 1707
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1748 GCAGAGAGGAACAAGGGGCTGAATTTGGTCAGCTTTTTCAGAAAGGATGGAGGATAT 1807
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1708 CTCGCATCTTATACAAAGAACCGGATGTTGGTGGGCAAGATGATGCGGACATTTCAAC 1767
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1808 CTTCTGGGTATTTATGAAGAAGCTGAATGGTGTAGCTGTGCTGATGTGCTTACCTTGAC 1867
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1768 GAATATGTGGAATAATGCCAAAGTGTCCATAGCACTTTCGACAGCTCGTACTAACTCAATC 1827
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1868 GAGTACATAAAGATGGAATCACATCTATCGGCCAACGTATATCTTCTGTTGAGTGGAGTG 1927
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1828 TTTTTCAC---TGGAGAAATTAATCTCTGATTAATTTTACAGCAAGTAGACCTTCCGTCC 1884
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1928 TTGATAAATGGATGGGCAACTCTCTTTCGCAAGAGGCATTTAGAAAAGTAGATTTATCCAGGA 1987
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1885 AAATTTCTGC-----ATCTTGTGCTTTTGAAGTGGAGCTAATCAATGACACCAAGACT 1938
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1988 AGACGTGTTCTCACAGAGCTGAATAGCTCAATTTCCCGCTCGCGGATGACACGAAGACA 2047
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1939 TACCAGCGGAGAGAAACCGTGGTGAATTTGGTTTTCCAGGGTACAGTGTCTACATGAGGAA 1998
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2048 TATAAAGCTGAGAGAGCTCGTGGAGAAATTTGGCTCCAGCATTTGATGTTACATGAAAGAC 2107
Db  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY	1999	AATCCGGAGTGCACAGAGGAAGAGCTCTAAGTCATGTTTATGGTATCATCGCAACGCA	2058
Db	2108	CATCCTGAATGTACAGAGGAAGAGGCTCTCGATCACAATCTATAGCATTCCTGGAGCCGGCG	2167
QY	2059	CTGAGGAAATGGAATTCGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA	2118
Db	2168	GTGAGGAACTGACAGAGAGTTCTGAGCCG---ACGAGTCCCATTCGCTGCAAG	2224
QY	2119	AGACTGCTGTTTCAACTGCAAGAGTGCAGCTGTTTATATGTAACAGATGGCTTT	2178
Db	2225	AAGATGCTTTTCGAGGAGACAAGAGTGACGATGGTGATATTCAAGGATGGAGATGGAATC	2284
QY	2179	GGTATCTGACAAAGAGATGAAGACCATGTGACCGCAACTCTTTTCGATCCTGTGCG	2237
Db	2285	GGTGTTCCTCAATTAAGAAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCCCTGCC	2343
RESULT 14			
US-09-234-393-39			
; Sequence 39, Application US/09234393A			
; Patent No. 6265639			
; GENERAL INFORMATION:			
; APPLICANT: Croteau, Rodney B			
; APPLICANT: Crook, John E			
; APPLICANT: Bohman, Jorg			
; APPLICANT: Steele, Christopher L			
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)			
; TITLE OF INVENTION: AND METHODS OF USE			
; FILE REFERENCE: WSUR113345			
; CURRENT APPLICATION NUMBER: US/09/234,393A			
; CURRENT FILING DATE: 1999-01-20			
; EARLIER APPLICATION NUMBER: 60/072,204			
; EARLIER FILING DATE: 1998-01-22			
; NUMBER OF SEQ ID NOS: 55			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 39			
; LENGTH: 2525			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence:			
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding			
; OTHER INFORMATION: E-alpha-bisabolene synthase			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1)..(2451)			
US-09-234-393-39			
Query Match			
Best Local Similarity 32.3%; Score 723.4; DB 3; Length 2525;			
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;			
QY	31	AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCGTCATTAATGACGGCTCTGTG	90
Db	238	ACTCCATCTGCTTATGACACAGCATGGGTAGCGAGGTGCCCGCCATTGATGGCTCTGCT	297
QY	91	GCACCCCAATTTCCCAACCGTTCAATCGATTCTGAACAATCAACTGCGCAGATGGCTCG	150
Db	298	CGCCCGCAATTTCCCAACAGTTGACTGGATTTTGAANAACCAAGTTAANAAGATGGTTCA	357
QY	151	TGGGGTGGAGAGTGCAATTTTCTGGGGTATGACAGAGTTTAAACACTCTCGCCTGCCTC	210
Db	358	TGGGGAATTCAGTCCCACTTTCTGCTGTCGACCGTCTTCTGGCACTCTTTCTGTGTT	417
QY	211	CTCAGTCTCAAAATATGGAATAGGGCGACATTCAGATGTCAGAAAGGGTTGATTTG	270
Db	418	CTTGCTCTCTTAAATGGAACGTTGGGGATCTCAAGTAGTAGAGCAGGGAATTTGAATTCAT	477
QY	271	AGAAACACATGGAAGAAATGAGGACGAGCTGACAAATCACAG---GCCAAGTGGAATC	327
Db	478	AAGAGCAATCTGGAATCTAGTAAAGGATGAACCGATCAAGATAGCTTGGTAACAGACTTT	537
QY	328	GAGTCTGCTTTCTCTGCAATGTTTAGATGAAGCAAAAAGCTTGGATTTGGATCTTCCTTTAT	387
Db	538	GAGATCATATTTCTCTTCTCTGTTAAGAGAAGCTCAATCTCTGCGCTCGGACTTCCCTAC	597
QY	388	CACCTCCTTTTATCTCCCAATCCACCAAAAGCGCCAGAAAAGCTTCAAAAGATTTCC	447
Db	598	GACCTGCTTTATATACATCTGTTGACAGCTAAACCGCAGGAAAGATTTAGCAAAAATTTCA	657
QY	448	CTCAATGCTTCTTCAATACCATCAGACGGCTGTGCTCTACTCTCTGGAGGGTTTGAAGAT	507
Db	658	AGGAGGAAATTTATGGGTTCCGTGCCATTTGTTGTTATCTTTAGAGGGAATACAAGAT	717
QY	508	GTGCTGAGCTGGCAAGAGATCACAAATCTTCAATCAAGAGACGGATCATTTTAAAGCTCC	567
Db	718	ATAGTTGAATGGGAACGAATAATGGAAGTTCAAAAGTCAGAGTGGGTCTTCTTAAAGCTCA	777
QY	568	CTGCACTACTGCTTGTGCTTCTCATGCACACTCAAAACAAACAGATGCTCCACATTTCTC	627
Db	778	CCTGCTTCTACTGCTGCTTTTCTGACACAGAGACGGGAATGCTTGAATTTCTTG	837
QY	628	AACCTTGTGCTCAGCAAAATTTGGGACTACCTTCTTCCCATTTACCCACTTGAATCTATTT	687
Db	838	AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGTATCTCTGTGATCTGCTG	897
QY	688	GAACGCTCTGGGCTGTCGATACAGTTGAACGCTTGGGAATCGATCGCTATTTCAAGAA	747
Db	898	GAACGCTTGTGATCGTAGATAATATGTACGCTTGGAAATCTATAGACACTTTGAAAAG	957
QY	748	GAAATCAAGAAATCTCTGGATTTAGCTTTATAGGTACTGGACCGCCGAAAGCGCTGGGA	807
Db	958	GAAATCAAGAAAGCTCTTGATTTATGTTTACAGGCATTTGAA---CGAAGAGGAATTTGG	1014
QY	808	TGGCAAGATGCAATCTTATCTGATGTCGATGACACTGCCCATGGCTTTTGAATCTCTG	867
Db	1015	TGGGCAGACTAAATCCCATAGACATCTTGAGACCACTGCTTTGGGATTTGGAATTCGTT	1074
QY	868	AGACTTCAATGATCAATGTTATCTTCAGATGTTCTGGAGAAATTTCAAGACACGAGAA	927
Db	1075	CGGCTGCATAGGTACAATGTATCTCCGACATTTTGTGACAACTTCAAAAGTGCATATGG	1134
QY	928	GACTTCTTTTGTGCTTGGCGTCAAAACGCAAAATGGTGTGACCGCAATCTTAACTTTAT	987
Db	1135	AAATTCATTTGCTCCACCGTCAATTCACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1194
QY	988	AGATGTTCAAGAATGTTTTCGGGAGAAAAGATAATGGAAGAGCTTAAGACCTTCACT	1047
Db	1195	AGAGCTTCCAGCTCGCATTTCCCGAGAAAACATTTCTTGATGAAGCTAAAAGCTTCGCT	1254
QY	1048	ACAAATCATCTCCAAATGCTCTTGCRAAAAACACGCAATTTGATAGTGGGCTGTCAAG	1107
Db	1255	ACTNAATATTTGAGAGAAGCTCTTGAGAAAAGTGAGACTTCCAGTGCATGGAAACACAA	1314
QY	1108	AAGGATCTTCTCGAGAGGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAGTATGCCA	1167
Db	1315	CAAACTGAGCCAAAGAGATCAATAACGCGCTGAAGACTTCTTGGCATGCGAGTGTCCG	1374
QY	1168	AGATTGGAGGCAAGATTTACATAGAGCAATTTGGATCAAAATGATGCTCTGGCTGGGGA	1227
Db	1375	AGAGTGGAGCAAGAGATGCTGCAAGTGTATCGCCCAAGATTTGCAACCATAGCAAAA	1434
QY	1228	ACTGTGTATAAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGCCAAATTTGG	1287
Db	1435	TGGTTTACAGCTACCTTACGTGAACATGAAGAATTTTATAGAGCTGGGAAAATTAGAT	1494
QY	1288	TTCAATATGGTGCAGGCTTTACACCAAAAGGAGACTCAACACATTTGTACGCTGGTGAGA	1347
Db	1495	TTCAACATTTATCCAGTCCATCCACCAAGAGAAATGAAGAATGTATACAGCTGGTTTGA	1554
QY	1348	GAATCGGATTCATGATCTTACATTTACATCCCGCAGCGCTGTGGAAATGTATTTCTCA	1407
Db	1555	GATTCGGGGTGGCACTATTTCACCTTCGCTCGGAGAGGCGCTGGAAATTTCTACTTTTA	1614
QY	1408	GTGGCGTTAGTATGTTTGAAGCAGAAATTCGCTGTTGTAGAAATTCCTATGCAAGACT	1467

Db 1615 GTAGCGGGGACCTATGAAACCCAGATGCGAAATGCAAGTTCCTCTTTACAAAGTG 1674  
Qy 1468 TCTTGCTCCAGCTAFTCTAGACGATCTTTAGCACCCACCGGATCTCTGAGTATCTT 1527  
Db 1675 GCATGCTTGACAGCTGTTCTGGACGATATGATGACACTTATGGAACCTTAGATGATTG 1734  
Qy 1528 AATATTGTTCTGACAGGCTCCAGATGGGATATCTCTGTCTGATGATGATGATGATGAT 1587  
Db 1735 AAGCTATTCTAGCTGAGGCTGGAGAGATGGGACCTCTCTCTTTACAGAAAACCTTCCAGAC 1794  
Qy 1588 AATCAGTTGAAGTTGCTTCTTAGGGCTGTACACACAGTGAATGATTTGGAAGAAT 1647  
Db 1795 TAT---ATGAACTATGTTACCAATCTATATGACATATGTTACAGAGTGGCTTGGAG 1851  
Qy 1648 GGAATCAAGGAAACAGCCGCTGATGCTGGCTATCTTCGAAAAGTATGGAGGGCTTG 1707  
Db 1852 GCAGAGAGGAAACAGGGGCTGAATGGTCACTCTTTTTCAGAAAAGGATGGAGGATTAT 1911  
Qy 1708 CTGCAATCGTATACCAAGAGCGGATGCTGGGAGCAAAAGTATGTCGCAATTCAC 1767  
Db 1912 CTTCTGGGTTATTTATGAAGAAGCTGAATGGTTAGCTGCTGAGTATGTCCTACCTTGGAC 1971  
Qy 1768 GAATATGTGGAATATCCAAATGTCATAGCACTTCGACACTGCTACTAACTCAATC 1827  
Db 1972 GAGTACATAAAGAAATGGAATCAATCTATCGGCAAGGCAATTAGAGAAAGTATGAGTG 2031  
Qy 1828 TTTTTCAC---TGGAGAAATTAATCTCTGATTAATTTACAGCAAGTAGACCTTCGGTCC 1884  
Db 2032 TTGATAATGATGGCAACTCTTTTCGCAAGAGGCAATTAGAGAAAGTATGAGTGAGTG 2091  
Qy 1885 AATTTCTGC-----ATCTTGCTGCTTTGAGTGGAGCACTAATCAATGACACCAAGCT 1938  
Db 2092 AGACGTGTTCTACAGAGCTGAATAGCTCAATTTCCCGCTGGCGGATGACAGCAAGACA 2151  
Qy 1939 TACCAGCCGAGAGAAACCGTGTGAATTTGTTTCCAGCGTACAGTGTACATGAGGAA 1998  
Db 2152 TATAAGCTGAGAAGCTCGTGAGAAATGGCTGCAAGCTTGAATGTTTACATGAAGAC 2211  
Qy 1999 AATCCGAGTGACAGAGGAAGCTCTAAGTCAATGTTTATGATATCATCGCAACGCA 2058  
Db 2212 CATCTGAAATGACAGAGGAAGGCTCTCGATCACTATAGACTTCTGGAGCCGGC 2271  
Qy 2059 CTGAAGAAATGAATTTGGAGTTGGCCAAACCCAGGAGCAATGCCCAATGTTGTGTAGA 2118  
Db 2272 GTGAAGAACTGACAGAGAGTTTCTGAAGCCG---ACGACGCTCCATTCGCTGCAAG 2328  
Qy 2119 AGACTGCTGTTCAACACTGCAAGAGTATGATGATGTTTATATGATGATGATGATGAT 2178  
Db 2329 AGATGCTTTTTCAGAGGACAAAGAGTGACGATGGTGATATTCAGGATGAGATGATTC 2388  
Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGACGCGCAACTCTTTTCGATCCTGTGGC 2237  
Db 2389 GGTGTTTCCAAATAGAGTCAAGATCATATCAAGAGTGTCTCAATGAAACCGCTGCC 2447

## RESULT 15

US-09-865-171-39  
; Sequence 39, Application US/09865171  
; Patent No. 6451576  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B  
; APPLICANT: Crock, John E  
; APPLICANT: Bohlmann, Jorg  
; APPLICANT: Steele, Christopher L  
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS),  
; FILE OF INVENTION: AND METHODS OF USE  
; FILE REFERENCE: WSUR117468  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 09/234,393  
; PRIOR FILING DATE: 1999-01-20  
; PRIOR APPLICATION NUMBER: 60/072,204

; PRIOR FILING DATE: 1998-01-22  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 39  
; LENGTH: 2525  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: computer-generated nucleic acid sequence encoding  
; OTHER INFORMATION: E-alpha-bisabolene synthase  
; NAME/KEY: CDS  
; LOCATION: (1)..(2451)  
; US-09-865-171-39

Query Match 32.3%; Score 723.4; DB 4; Length 2525;  
Best Local Similarity 59.8%; Pred. No. 3.9e-22;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

Qy 31 AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCGTCATATTGACGGCTCTGGT 90  
Db 238 ACTCCATCTGCTTATGACACAGCATGGGTAGCGGGTGCCTCCCATTTGATGGCTCTGCT 297  
Qy 91 GCACCCCAATTTCCCAACAGCTTCAATGGATTTCTGAACATCAACTGCCAGATGGCTCG 150  
Db 298 CGCCCGCAATTTCCCAACAGCTTGAATTTGAAAAACCACTTAAAAAGATGGTTCA 357  
Qy 151 TGGGGTGGAGGATGATCTTTCTGGGGTATGACAGAGATTTTAAACACTCTCGCTCGCTC 210  
Db 358 TGGGGAATTCAGTCCCACTTTCTGCTGTCGACCGTCTTCTTCCCACTCTTCTTGTGTT 417  
Qy 211 CTCACCTCTCAAAATATGGAATTAAGGGGACATTCAGAGTCAGAAAAGGGTGTGATTTGTG 270  
Db 418 CTTGTGCTCTCTTAAATATGGAACGTTGGGGATCTGCAAGTAGAGCAGGAAATTTGAATTCATA 477  
Qy 271 AGAACAACATGGAAGAAATGAAGGACGAGCTGCAATCACAG---GCCAAGTGAATTC 327  
Db 478 AGAGCAATCTGGAATCTAGTAAAGGATGAACCGCATCAAGTAGCTTGGTGAACAGACTTT 537  
Qy 328 GAGTCTGCTGTTCTCTCAATGTTAGATGAAGCAAAAAGCTTGGGATTTGGAATCTTCTTTAT 387  
Db 538 GAGATCATATTTCTCTCTCTGTTAAGAGAGCTCAATCTCTGCGCTCGGACTTCCCTAC 597  
Qy 388 CACCTCCCTTTCTATCTCCCAATCCACCAAGCGGACAGAAAAGCTTCAAAAAGATTTCC 447  
Db 598 GACCTGCTTATATATATCTGTTGCAGACTTAAACCGGAGGAAATAGCAAAACTTTCA 657  
Qy 448 CTCATGTTCTTTCATACCATCAGAGCGGTTGCTCTACTCTCTGGAGGGTTCGAGAT 507  
Db 658 AGGAGGAAATTTATCGGTTCCGTCGCCATTTGTTATTTCTTTAGAGGGAATCAAGAT 717  
Qy 508 GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGCGGATCATTTTAAAGTCC 567  
Db 718 ATAGTTGAATGGGAACGAATATGGAAGTTCAAGTCAGGATGGGTTCTTTTAAGTCA 777  
Qy 568 CTGCACTACTGCTGTTGTCTTTCATGCACTCAAAACCAAAAGATGCTCCCACTTTCTC 627  
Db 778 CTTGCTTCTACTGCTGCTGTTTTCATGCAACAGGAGAGCGGAAATGCTTGAATCTTG 837  
Qy 628 AACTTCGCTCAGCAAAATTTGGCGACTAGTTCTTCCATTTACCCACTTTGATCTATTT 687  
Db 838 AACAGTGTGATGATCAAGTTTGGAAATTTTGTCTTCCCTGCTGATCCTGTGGATCTGCTG 897  
Qy 688 GAACGCTCTGGGCTGCTGATACAGTTGAACGCTTGGGAATCGATCGTATTTTCAGAAA 747  
Db 898 GAACGCTCTGTTGATCGTATATATTTGACGCTTGGATCTTATAGACACTTTTGAAG 957  
Qy 748 GAAATCAAGAAATCTCTGATTAAGTTTATAGTACTGGAAGCCGCAAGAGCGGTGGGA 807  
Db 958 GAAATCAAGGAAGCTCTTGAATTTATAGGCAATTTGAA---CGAAGAGGAATTTGGG 1014  
Qy 808 TGGGCAAGATGCATCTTATCTGATGTCGATGACACTGCCCATGGTCTTAGAATCTG 867



Thu Aug 26 08:27:53 2004

Db 1015 TGGGCGAGACTAAATCCCATAGCAGATCTTGAGACCACTGCTTTTGGGATTTGCAATGCTT 1074  
 Qy 868 AGACTTCATGGATACAAATGATCTTCCAGATGTTCTGGAGAAATTTCCAGACGAGAAAGGA 527  
 Db 1075 CGGCTGATAGGTACAAATGATCTTCCAGCAATTTTGGACAACTTTTCAAGATGCCATGGG 1134  
 Qy 928 GACTCTCTTTGCTTTGCGGCTCAAGCGCAAAATTTGGTGTGACCGGATTAATCTTAACTTTAT 987  
 Db 1135 AAATTCATTTGCTCGACCGGTCAATTCACAAAGATGTAGCAAGCATGCTGAACTCTTTAT 1194  
 Qy 988 AGATGTTCAAGATGATGTTTTCGGGAGAAAGATTAATGGAAGCAAGCTTAAGACTTCACT 1047  
 Db 1195 AGAGCTTCCAGCTCGCATTTCCGGGAGAAACATTTCTTGATGAAGCTAAAGACTTCGCT 1254  
 Qy 1048 ACAATCATCTCCAAATGCTTTGCCCAAAACCAACGCAATTTGATAGTGGGCTGTCAAG 1107  
 Db 1255 ACTAATATTTGAGAGAGCTCTTTGAGAAAGTGAGACTTCCAGTGCAATGGAACAAACAA 1314  
 Qy 1108 AAGGATCTTCTCGAGAGGTGGAGTATGCTTATAAAGTATCCGGTGGCATAGAATGCCA 1167  
 Db 1315 CAAACCTGAGCCCAAGAGATCMAATACGCGCTGAAGACTTCTTGGCATGCCAGTGTTCG 1374  
 Qy 1168 AGATTGAGGCAAGAGTTACATAGAGCAATTTGGATCAAAATGATGCTCGCTGGGGAAG 1227  
 Db 1375 AGAGTGAAGCAAGAGATGATCTCAAGTGTATGCCAGATTTATGACGATAGCAAA 1434  
 Qy 1228 ACTGTGATAGATGCTATATGTGAGCAACGAAATAATTTGGAGCTGGCCAAATTTGGAC 1287  
 Db 1435 TGGGTTTACAGCTACCTCAGTGAACAAATGAAAGTTTTTGGAGCTGGGAAATTAGAT 1494  
 Qy 1288 TTCAATATGGTGCAGGCTTACACCAAAAGAGACTCAACATGTCAGCTGGTGGAGA 1347  
 Db 1495 TTCAACATTTCCAGTCCATCCACCAAGAGAAATGAAGATGTACAGCTGGTTTGA 1554  
 Qy 1348 GAATCGGATTCATGATCTTACATTCACCCGAGCGGCTGTGGAAATGATTTCTCA 1407  
 Db 1555 GATTGGGGTTGCCACTATTACCTTCGCTGGGAGAGCGCTGGAACTTCTACTTCTTA 1614  
 Qy 1408 GTGGCGTTAGTATGTTTGGCCAGAAATTCGCTGTTGTAGAAATTCCTATGCCAGACT 1467  
 Db 1615 GTAGCGGGGACCTATGAACCCAGTATGCCAAATGCAGGTTCCTCTTTACAAAGTG 1674  
 Qy 1468 TCTTGCCTCGAGTTATTTCTAGACCATTTTACGACACCCAGGATCTCTGGATGATCTT 1527  
 Db 1675 GCATGCTTGCAGACTGTTCTGAGCAGATATGATGACATTTATGGAACCTAGATGAATG 1734  
 Qy 1528 AAATTTGTTCTCTGAAGCGTCCGAAGATGGGATATCTCTGTGCTGGATAGCGTTGGGAT 1587  
 Db 1735 AAGCTATTCACTGAGGCTGTGAGAGATGGGACCTCTCCTTTACAGAAACCTTCCAGAC 1794  
 Qy 1588 AATCAGTTGAAAGTTTGTCTTCTAGGGCTGTACACACAGTGAATGATTTGGAAAGAT 1647  
 Db 1795 TAT---ATGAAACTATGTTACCAAAATCTATTATGACATAGTTCAAGGTGCTTGGAG 1851  
 Qy 1648 GGACTCAAGGAAACAGCGCTGATGCTGGGCTATCTTTGAAAGATATGGAGGGCTTG 1707  
 Db 1852 GCAGAGAGGAAACAGGGGGTGAATTTGGTCAGCTTTTTCAGAAAGGATGGGAGGATAT 1911  
 Qy 1708 CTGCGCATCGTATACCAAGAGCCGAATGTTGGGAGCAAGATGATGTCGGCATTTCAAC 1767  
 Db 1912 CTTCTGGGTTATATGAAGAAGCTGAATGGTTAGCTGCTGATGTTGCTTACCTTGGAC 1971  
 Qy 1768 GAATATGTGGAATGCAAGATGTCATAGACATTTGGAGACTTGGACAGTCTGTAACCTCAATC 1827  
 Db 1972 GAGTACATAAAGAAATGGAATCAATCTATCGGCCAAGTATATCTTGTGAGTGGAGTG 2031  
 Qy 1828 TTTTTCAC---TGGAGAAATTAATCTTCTGATTTACATTTTACGCAAGTAGACCTTCGGTCC 1884  
 Db 2032 TTGATATGATGGCAACTCTTTTCGCAAGAGGCAATTAGAGAAAGTAGATTTATCCAGGA 2091  
 Qy 1885 AAATTTCTGC-----ATCTTGTGCTTTTCTGAGGACTTAATCATATGACACCAAGACT 1938  
 Db 2092 AGACGTGTTCTCAGAGAGCTGAATAGCCTCATTTTCCCGCTTGGCGGATGACACGAAGACA 2151

Qy 1939 TACCAGCCGAGAGAAACCGTGTGAATTTGGTTCAGCGTACAGTGTCTACATGAGGAA 1998  
 Db 2152 TATAAAGCTGGAAGGCTCGTGAGAAATTTGGGCTCCAGCATTCGAATGTTTACATGAAGAC 2211  
 Qy 1999 AATCCGAGTGCACAGAGAGAAAGCTCTTAAGTCAATGTTTATGTTATCATCGACAACGCA 2058  
 Db 2212 CATCTGATGTACAGAGAGAGGCTCTCGATCACATCTATAGCATTTCTGGAGCCGGCG 2271  
 Qy 2059 CTGAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCGAGCAATGCCCATTTGTGTGAGA 2118  
 Db 2272 GTGAAGGAATCTGAAGAGAGATTTCTGAAGCCCG---ACGACGTCCCATTTCCGCTGCAAG 2328  
 Qy 2119 AGACTGCTGTTCAACTGCAAGAGTGTGAGCTGTTTATATGTACAGAGATGGCTTTT 2178  
 Db 2329 AAGATGCTTTTCAGGAGACAGAGTGACGATGGTGATATTCAGGATGGAGATGGATTC 2388  
 Qy 2179 GGTATCTCTGACAAAGAGATGAAAGACCATGTGAGCGGAACCTCTTTTCGATCCTGTGGC 2237  
 Db 2389 GGTGTTTCCAAATTTAGAAAGTCAAAGATCATATCAAAGAGTGTCTCTCATTTGAACCGCTGCC 2447

Search completed: August 24, 2004, 21:21:05  
 Job time : 153.974 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2241	100.0	2241	14	US-10-041-007-38	Sequence 38, App
2	2241	100.0	2388	14	US-10-041-007-36	Sequence 36, App
3	2241	100.0	2445	14	US-10-041-007-34	Sequence 34, App
4	2241	100.0	2622	14	US-10-041-007-32	Sequence 32, App
5	2241	100.0	2705	12	US-10-041-018-397	Sequence 397, App
6	2241	100.0	2705	14	US-10-041-007-1	Sequence 1, App1
7	1078	48.1	2861	9	US-09-887-586A-55	Sequence 55, App1
8	1078	48.1	2861	9	US-09-903-012-55	Sequence 55, App1
9	1078	48.1	2861	10	US-09-900-737-55	Sequence 55, App
10	1078	48.1	2861	12	US-10-041-018-365	Sequence 365, App
11	1078	48.1	2861	13	US-09-893-820-55	Sequence 55, App
12	1078	48.1	2861	14	US-10-041-007-3	Sequence 3, App1
13	723.4	32.3	2424	9	US-09-887-586A-45	Sequence 45, App1
14	723.4	32.3	2424	9	US-09-903-012-45	Sequence 45, App1

181 GACAGAGTTTAAACACTCTCGCTCGCTCCTCACTCTCAAAATATGGAATAGGCGGAC 240  
181 GACAGAGTTTAAACACTCTCGCTCGCTCCTCACTCTCAAAATATGGAATAGGCGGAC 240  
241 ATTCAAGTCAGAAAGGGGTGAGTTGTGAGAAAACACATGGAAGAAATGAGGACGAA 300  
241 ATTCAAGTCAGAAAGGGGTGAGTTGTGAGAAAACACATGGAAGAAATGAGGACGAA 300  
301 GCTGACAAATCACAGGCCAGTGGATTCGAGTCTGCTGCTTCTCTCAATGTTAGATGAAGCA 360  
301 GCTGACAAATCACAGGCCAGTGGATTCGAGTCTGCTGCTTCTCTCAATGTTAGATGAAGCA 360  
361 AAAAGCTTGGGATGGAATCTTCTTATPACCTCCCTTTTCACTCCCAATCCCAAAAG 420  
361 AAAAGCTTGGGATGGAATCTTCTTATPACCTCCCTTTTCACTCCCAATCCCAAAAG 420  
421 CGCCAGAAAAGCTTCAAAAGATCCCTCAATGTTCTTCAATACCATCAGACGGCGTTG 480  
421 CGCCAGAAAAGCTTCAAAAGATCCCTCAATGTTCTTCAATACCATCAGACGGCGTTG 480  
481 CTCTACTCTCTGAGAGGTTTGAAGATGTGTGACTGGCAAGAGATCAAAATCTTCAA 540  
481 CTCTACTCTCTGAGAGGTTTGAAGATGTGTGACTGGCAAGAGATCAAAATCTTCAA 540  
541 TCAAGAGACGGATCAATTTTAAAGCTCCCTGCTCTACTGCTTCTTCAATGCTTCA 600  
541 TCAAGAGACGGATCAATTTTAAAGCTCCCTGCTCTACTGCTTCTTCAATGCTTCA 600  
601 CAAACCAACCGATCCCTCACTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTACGTT 660  
601 CAAACCAACCGATCCCTCACTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTACGTT 660  
661 CTTTGGCAATACCACTTGAATCTATTTGAAAGCGCTCTGGGCTGCTGATGTAAGG 720  
661 CTTTGGCAATACCACTTGAATCTATTTGAAAGCGCTCTGGGCTGCTGATGTAAGG 720  
721 TTGGGATCGATCGTATTTCAAGAGAAATCAAGAAATCTCTGATTAAGTATATAGG 780  
721 TTGGGATCGATCGTATTTCAAGAGAAATCAAGAAATCTCTGATTAAGTATATAGG 780  
781 TACTGGGACGCGCAAGAGGCTGGGATGGGCAAGATGCAATCTTCTGATGTCGAT 840  
781 TACTGGGACGCGCAAGAGGCTGGGATGGGCAAGATGCAATCTTCTGATGTCGAT 840  
841 GACATGCGCATGGGCTTTAGATCTGAGACTTCAATGATGATGATGATGATGATGAT 500  
841 GACATGCGCATGGGCTTTAGATCTGAGACTTCAATGATGATGATGATGATGATGAT 500  
901 CTGGAGAAATTTCAAGAGAGAGAGAGACTTCTTTTCTGCTTCTGCTTCTGCTTCT 960  
901 CTGGAGAAATTTCAAGAGAGAGAGAGACTTCTTTTCTGCTTCTGCTTCTGCTTCT 960  
961 GGTGTGACCGATTAATCTTAACTTTATAGATGTTCAAGATGTTTCTGGGAGAAAG 1020  
961 GGTGTGACCGATTAATCTTAACTTTATAGATGTTCAAGATGTTTCTGGGAGAAAG 1020  
1021 ATAAATGGAAGAGCTAAGACCTTCACTACAAATCACTCTCAAAATGCTTCTGCAAAAAC 1080  
1021 ATAAATGGAAGAGCTAAGACCTTCACTACAAATCACTCTCAAAATGCTTCTGCAAAAAC 1080  
1081 AACGCAATTTGATAAGTGGGCTGTCAAGAGGATCTCTCTGGAGAGGTTGATGCTATA 1140  
1081 AACGCAATTTGATAAGTGGGCTGTCAAGAGGATCTCTCTGGAGAGGTTGATGCTATA 1140  
1141 AAGTATCCGTGGCATAGAGTATGCCAAGATTCGAGCAAGAGTATACATAGAGCAATTT 1200  
1141 AAGTATCCGTGGCATAGAGTATGCCAAGATTCGAGCAAGAGTATACATAGAGCAATTT 1200  
1201 GGATCAATGATGCTGCTGGGAGAGACTGTTATGATGATGATGATGATGATGATGAT 1260  
1201 GGATCAATGATGCTGCTGGGAGAGACTGTTATGATGATGATGATGATGATGATGAT 1260  
1261 AAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGTTGGTGGAGGCTTACACCAAAAGGAG 1320

1261 AAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGTTGGTGGAGGCTTACACCAAAAGGAG 1320  
1321 ACTCAACACATTTGTAGCTGGTGGAGAGAAATCGGATTAATGATCTTCAATTCACCCGC 1380  
1321 ACTCAACACATTTGTAGCTGGTGGAGAGAAATCGGATTAATGATCTTCAATTCACCCGC 1380  
1381 CAGCGCCTGTGAAATGTAATTTCTCAGTGGCGGTTAGTATGTTGAGCCAGAAATTCGCT 1440  
1381 CAGCGCCTGTGAAATGTAATTTCTCAGTGGCGGTTAGTATGTTGAGCCAGAAATTCGCT 1440  
1441 GCTTGTAGAAATCCCTATGCAAGACATTTCTGCTGCTGCAATTTCTAGACGATCTTTAC 1500  
1441 GCTTGTAGAAATCCCTATGCAAGACATTTCTGCTGCTGCAATTTCTAGACGATCTTTAC 1500  
1501 GACACCCACGGATCTCTGGAATGATCTTAAATTTGTTCTCTGAAGCGTCCGAAAGTGGAT 1560  
1501 GACACCCACGGATCTCTGGAATGATCTTAAATTTGTTCTCTGAAGCGTCCGAAAGTGGAT 1560  
1561 ATCTCTGCTGATAGCTTTCGGGATTAATCAGTTGAAAGTTTCTCTCTAGGCTGTAC 1620  
1561 ATCTCTGCTGATAGCTTTCGGGATTAATCAGTTGAAAGTTTCTCTCTAGGCTGTAC 1620  
1621 AACACAGTGAATGGAATTTGGAAGATGGACTCAAGGAAACGAGCCGCTGATGCTGGGC 1680  
1621 AACACAGTGAATGGAATTTGGAAGATGGACTCAAGGAAACGAGCCGCTGATGCTGGGC 1680  
1681 TATCTTCTGAAAGTATGGAGGCTTCTGCTGCAATCTGATATCAAAAGAGCGGAAATGCTG 1740  
1681 TATCTTCTGAAAGTATGGAGGCTTCTGCTGCAATCTGATATCAAAAGAGCGGAAATGCTG 1740  
1741 GCAGCAAGATGTCGCGACATTCGATCGATCGATATCAAAAGAGCGGAAATGCTGCTG 1800  
1741 GCAGCAAGATGTCGCGACATTCGATCGATCGATATCAAAAGAGCGGAAATGCTGCTG 1800  
1801 CTTGCGAGCTGCTACTTAACTCAATCTTTTCTGCTGAGAAATTTACTTCTGATTTACTT 1860  
1801 CTTGCGAGCTGCTACTTAACTCAATCTTTTCTGCTGAGAAATTTACTTCTGATTTACTT 1860  
1861 TTACAGCAAGTAGACCTTTCGGTCCAAATTTCTGCAATCTTGTCTTTGACTGGAGACTA 1920  
1861 TTACAGCAAGTAGACCTTTCGGTCCAAATTTCTGCAATCTTGTCTTTGACTGGAGACTA 1920  
1921 ATCAATGACACCAAGACTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
1921 ATCAATGACACCAAGACTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
1981 CAGTGTCTATCAGAGGAAATCCGGAGTGCACAGAGGAAAGGCTCTAAGTCAATGTTTAT 2040  
1981 CAGTGTCTATCAGAGGAAATCCGGAGTGCACAGAGGAAAGGCTCTAAGTCAATGTTTAT 2040  
2041 GGTATCATCGACACCGCTGGAAGAAATTTGGAATTTGGAGTTGGCCAAACCCAGCGAGCAAT 2100  
2041 GGTATCATCGACACCGCTGGAAGAAATTTGGAATTTGGAGTTGGCCAAACCCAGCGAGCAAT 2100  
2101 GCCCAATTTCTGTTGAGAGAGCTGCTGTTTCAACACTCTGCAAGAGTGTAGAGCTGTTTAT 2160  
2101 GCCCAATTTCTGTTGAGAGAGCTGCTGTTTCAACACTCTGCAAGAGTGTAGAGCTGTTTAT 2160  
2161 ATGTACAGAGATGCTGTTGGTATCTCTGACAAAGAGATGAAAGCCATGTCAGCCGAAC 2220  
2161 ATGTACAGAGATGCTGTTGGTATCTCTGACAAAGAGATGAAAGCCATGTCAGCCGAAC 2220  
2221 CTTTTCGATCTGTCGCTAG 2241  
2221 CTTTTCGATCTGTCGCTAG 2241

RESULT 2  
US-10-041-007-36  
; Sequence 36, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:

; APPLICANT: Matsuda, Seiichi P.T. ; APPLICANT: Schepmann, Hala G. ; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthese ; FILE REFERENCE: P02081U1 ; CURRENT APPLICATION NUMBER: US/10/041.007 ; CURRENT FILING DATE: 2002-01-07 ; PRIOR APPLICATION NUMBER: US 60/259,881 ; PRIOR FILING DATE: 2001-01-05 ; NUMBER OF SEQ ID NOS: 41 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 36 ; LENGTH: 2388 ; TYPE: DNA ; ORGANISM: Ginkgo biloba ; US-10-041-007-38									
Query Match 100.0%; Score 2241; DB 14; Length 2388; Best Local Similarity 100.0%; Pred. No. 0; Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGTTTCAGTCCATGGGCGACGGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTG	50						
Db	148	ATGTTTCAGTCCATGGGCGACGGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTG	207						
QY	61	GCAGAAATTCGGTCAATTTGACGGCTCTGTGTCACCCCAATTTCCCAACGCTTCAATGG	120						
Db	208	GCAGAAATTCGGTCAATTTGACGGCTCTGTGTCACCCCAATTTCCCAACGCTTCAATGG	267						
QY	121	ATTCTGAACAATCAATGCCAGATGGCTCGTGGGGTGAGAGTGCATTTTCTGGCGTAT	180						
Db	268	ATTCTGAACAATCAATGCCAGATGGCTCGTGGGGTGAGAGTGCATTTTCTGGCGTAT	327						
QY	181	GACAGAGTTTAAACACTCTCGCTGCTCTCACTCTCAAAATATGGAATAGGGCGAC	240						
Db	328	GACAGAGTTTAAACACTCTCGCTGCTCTCACTCTCAAAATATGGAATAGGGCGAC	387						
QY	241	ATTCAAGTCAGAAAGGGTGTGTTGTGAGAAACACATGAGAAATGAGAGCGAA	300						
Db	388	ATTCAAGTCAGAAAGGGTGTGTTGTGAGAAACACATGAGAAATGAGAGCGAA	447						
QY	301	GCTGACAAATCACAGGCCAAGTGCAGGTGGTCTGCAATGTTAGATGAAGCA	360						
Db	448	GCTGACAAATCACAGGCCAAGTGCAGGTGGTCTGCAATGTTAGATGAAGCA	507						
QY	361	AAAAGCTTGGGATTTGATCTTCACTCACTCCCTTCACTCCCAATCCCAAAAG	420						
Db	508	AAAAGCTTGGGATTTGATCTTCACTCACTCCCTTCACTCCCAATCCCAAAAG	567						
QY	421	CGCCAGAAAGCTTCAAAAGATCCCTCAATGTTCTTCAACCAATCAGACGGCGTTG	480						
Db	568	CGCCAGAAAGCTTCAAAAGATCCCTCAATGTTCTTCAACCAATCAGACGGCGTTG	527						
QY	481	CTCTACTCTCTGGAGGTTTGCAGAGATGTGTGACCTGGCAAGAGATCACAAATCTTCAA	540						
Db	628	CTCTACTCTCTGGAGGTTTGCAGAGATGTGTGACCTGGCAAGAGATCACAAATCTTCAA	587						
QY	541	TCAAGAGACGGATCATTTTTAAGCTCCCTGCGATCTACTGCTTGTGTCTTCATGCACT	500						
Db	688	TCAAGAGACGGATCATTTTTAAGCTCCCTGCGATCTACTGCTTGTGTCTTCATGCACT	747						
QY	601	CAAAACAAACGATGCCTCCACTTCTCACTTGTGTCTGATGAGAAATTTGGCGACTAGTT	560						
Db	748	CAAAACAAACGATGCCTCCACTTCTCACTTGTGTCTGATGAGAAATTTGGCGACTAGTT	807						
QY	661	CTTTGCCATTACCACTTGTATTTTGAACGCTCTGGGCTGTGCATACAGTTGAACGC	720						
Db	808	CTTTGCCATTACCACTTGTATTTTGAACGCTCTGGGCTGTGCATACAGTTGAACGC	867						
QY	721	TTGGGAATCGATCGTATTTTCAAGAAAGAAATCAAGAAATCTCTGGAATACGTTTATAGG	780						
Db	868	TTGGGAATCGATCGTATTTTCAAGAAAGAAATCAAGAAATCTCTGGAATACGTTTATAGG	927						
QY	781	TACTGGACGCCGAAAGAGCGGTGGGATGGGCAAGATGCAATCCTATCTCTGATGTCGAT	940						
Db	928	TACTGGACGCCGAAAGAGCGGTGGGATGGGCAAGATGCAATCCTATCTCTGATGTCGAT	987						
QY	841	GACACTGCCATGGGTCTTGAATCCTGAGACTTTCATGGATACAAATGATCTTCAGATGTT	900						
Db	988	GACACTGCCATGGGTCTTGAATCCTGAGACTTTCATGGATACAAATGATCTTCAGATGTT	1047						
QY	901	CTGGAGAAATTTCAAGACGAGAAAGAGACTCTCTTTGCTTTGCGGCTCAACGCAATTT	960						
Db	1048	CTGGAGAAATTTCAAGACGAGAAAGAGACTCTCTTTTGGCTTTGCGGCTCAACGCAATTT	1107						
QY	961	GGTGTGACCGATAATCTTAACTTTTATAGATGTTTCAAGATGATGTTTTCGGGAGAAAG	1020						
Db	1108	GGTGTGACCGATAATCTTAACTTTTATAGATGTTTCAAGATGATGTTTTCGGGAGAAAG	1167						
QY	1021	ATAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGGCCAAAC	1080						
Db	1168	ATAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGGCCAAAC	1227						
QY	1081	AACGCAATTTGATTAAGTGGGCTGTCAAGAGGATCTTCTCGAGAGGTGGAGTATGCTATA	1140						
Db	1228	AACGCAATTTGATTAAGTGGGCTGTCAAGAGGATCTTCTCGAGAGGTGGAGTATGCTATA	1287						
QY	1141	AAGTATCCGTGGCATAGAGTATGCCAAGATTCGAGGCAAGAGTTACATAGAGCAATTT	1200						
Db	1288	AAGTATCCGTGGCATAGAGTATGCCAAGATTCGAGGCAAGAGTTACATAGAGCAATTT	1347						
QY	1201	GGATCAAAATGATCTCTGGCTGGGGAAGACTGTGTATAAGATGCTTATGTGAGCAACGAA	1260						
Db	1348	GGATCAAAATGATCTCTGGCTGGGGAAGACTGTGTATAAGATGCTTATGTGAGCAACGAA	1407						
QY	1261	AAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGTTGAGCCGCTTACACCAAGAGAG	1320						
Db	1408	AAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGTTGAGCCGCTTACACCAAGAGAG	1467						
QY	1321	ACTCAACACATTTGCTGAGCTGGTGAGAGATCGGATTCATATGATCTTACATTCACCCGC	1380						
Db	1468	ACTCAACACATTTGCTGAGCTGGTGAGAGATCGGATTCATATGATCTTACATTCACCCGC	1527						
QY	1381	CAGCGGCTGTGGAATGTAATTTCTCAGTGGCGGTTAGTATGTTTGAAGCCAGAAATTCGCT	1440						
Db	1528	CAGCGGCTGTGGAATGTAATTTCTCAGTGGCGGTTAGTATGTTTGAAGCCAGAAATTCGCT	1587						
QY	1441	GCTTGTAGAAATTCCTATGTCACAAAGACTTCTTGGCTCGCAGTATTTCTAGAGATCTTTAC	1500						
Db	1588	GCTTGTAGAAATTCCTATGTCACAAAGACTTCTTGGCTCGCAGTATTTCTAGAGATCTTTAC	1647						
QY	1501	GACACCCAGGATCTCTGGATGATCTTAAATGCTCTGCAAGCGGTCCGAAAGATGGAT	1560						
Db	1648	GACACCCAGGATCTCTGGATGATCTTAAATGCTCTGCAAGCGGTCCGAAAGATGGAT	1707						
QY	1561	ATCTCTGTGTGATAGCGTTCCGGATTAATCAGTTGAAAGTTTGTCTTCTAGGGCTGTAC	1620						
Db	1708	ATCTCTGTGTGATAGCGTTCCGGATTAATCAGTTGAAAGTTTGTCTTCTAGGGCTGTAC	1767						
QY	1621	AACAAGTGAATGGAATTTGAAAGATGGAATCAAGAAACAGGCGGTGATGCTGGGC	1680						
Db	1768	AACAAGTGAATGGAATTTGAAAGATGGAATCAAGAAACAGGCGGTGATGCTGGGC	1827						
QY	1681	TATCTTGGAAAGATGAGGAGGCTTGGCTGCGATCGTATACCAAGAGCGGATGCTGCG	1740						
Db	1828	TATCTTGGAAAGATGAGGAGGCTTGGCTGCGATCGTATACCAAGAGCGGATGCTGCG	1887						
QY	1741	GCAGCAAGATGATGTCGCAATTTCAACGAATATGTGGAATATGCAAAAGTGTCCATAGCA	1800						
Db	1888	GCAGCAAGATGATGTCGCAATTTCAACGAATATGTGGAATATGCAAAAGTGTCCATAGCA	1947						
QY	1801	CTTGGACAGTCTGATCAAACTCAATCTTTTTCATCTGGAGAAATTAATCTCTGATTAATT	1860						
Db	1948	CTTGGACAGTCTGATCAAACTCAATCTTTTTCATCTGGAGAAATTAATCTCTGATTAATT	2007						
QY	1861	TTACAGCAAGTACACCTTGGTCCAAATTTCTGCAATCTTGTGCTTTGACTGGACGACTA	1920						

Db	2008	TTACAGCAAGTAGACCTTCGGTCCAAATTTCTGCATCTTGTGTCTTTGACTGGACGACTA	2067
Qy	1921	ATCAATGACACCAAGACTTACAGAGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA	1980
Db	2068	ATCAATGACACCAAGACTTACAGAGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA	2127
Qy	1981	CAGTGTACATGAGGAAATCCGGAGTGACAGAGGAAGAAGCTCTAAAGTCATGTTAT	2040
Db	2128	CAGTGTACATGAGGAAATCCGGAGTGACAGAGGAAGAAGCTCTAAAGTCATGTTAT	2187
Qy	2041	GGTATCATGACAAACGACCTGAAGGAATTCGAATTTGGAGTTGGCCACCCAGCGAGCAAT	2100
Db	2188	GGTATCATGACAAACGACCTGAAGGAATTCGAATTTGGAGTTGGCCACCCAGCGAGCAAT	2247
Qy	2101	GGCCCATGTTGTGAGAAAGACTGCTCTCAACACTGCAAGAGTGATGACAGCTGTTTTAT	2160
Db	2248	GGCCCATGTTGTGAGAAAGACTGCTCTCAACACTGCAAGAGTGATGACAGCTGTTTTAT	2307
Qy	2161	ATGTPACAGAGATGGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTCCAGCGAAT	2220
Db	2308	ATGTPACAGAGATGGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTCCAGCGAAT	2367
Qy	2221	CTTTTCGATCCTGTGGCGTAG	2241
Db	2368	CTTTTCGATCCTGTGGCGTAG	2388
RESULT 3			
US-10-041-007-34			
; Sequence 34, Application US/10041007			
; Publication No. US20020164736A1			
; GENERAL INFORMATION:			
; APPLICANT: Matsuda, Seiichi P.T.			
; APPLICANT: Schepmann, Hala G			
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase			
; FILE REFERENCE: P02081US1			
; CURRENT APPLICATION NUMBER: US/10/041,007			
; CURRENT FILING DATE: 2002-01-07			
; PRIOR APPLICATION NUMBER: US 60/259,881			
; PRIOR FILING DATE: 2001-01-05			
; NUMBER OF SEQ ID NOS: 41			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 34			
; LENGTH: 2445			
; TYPE: DNA			
; ORGANISM: Ginkgo biloba			
US-10-041-007-34			
Query Match			
Best Local Similarity 100.0%; Score 2241; DB 14; Length 2445;			
Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	ATGTTTCAGTCCATGGGCGACGGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTG	60
Db	205	ATGTTTCAGTCCATGGGCGACGGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTG	264
Qy	61	GCAGAAATTCCTCAATTCAGCGCTCTGTGCAACCCCAATTTCCCAACCGTTTCAATGG	120
Db	265	GCAGAAATTCCTCAATTCAGCGCTCTGTGCAACCCCAATTTCCCAACCGTTTCAATGG	324
Qy	121	ATTCTGAACAATCACTGCCAGATGGCTCGTGGGGTGAGAGTGCAATTTTCTGGCGTAT	180
Db	325	ATTCTGAACAATCACTGCCAGATGGCTCGTGGGGTGAGAGTGCAATTTTCTGGCGTAT	384
Qy	181	GACAGAGTTTAAACACTCTCGCTCGCTCTCACTCAATAATGGAATAGAGGCGAC	240
Db	385	GACAGAGTTTAAACACTCTCGCTCGCTCTCACTCAATAATGGAATAGAGGCGAC	444
Qy	241	ATTCAAGTGAGAAAGGGGTTGAGTTTGTGAGAAAAACACATGGAAGAAATGAAGGACGAA	300
Db	445	ATTCAAGTGAGAAAGGGGTTGAGTTTGTGAGAAAAACACATGGAAGAAATGAAGGACGAA	504
Qy	301	GCTGACAATCACAGGCCCAAGTGGATTCGAGTCTGTGTTTCTTCTGCAATGTTAGATGAAGCA	360

Db	505	GCTGCAATCACAGGCCAAGTGGATTCGAGTCTGTTCTTCTGCAATGTTAGATGAAGCA	564
Qy	361	AAAAGCTTGGAGTTGATCTTCTTATCACTCCCTTTTATCTCCCAATCCACCAAAAG	420
Db	565	AAAAGCTTGGAGTTGATCTTCTTATCACTCCCTTTTATCTCCCAATCCACCAAAAG	624
Qy	421	CGCCAGAAAAGCTTCAAAAGATTCCCTCAATGTTCTTCAATACCATCAGACGGGTTG	480
Db	625	CGCCAGAAAAGCTTCAAAAGATTCCCTCAATGTTCTTCAATACCATCAGACGGGTTG	684
Qy	481	CTCTACTCTCTGAGAGGTTTTCGAAGATGTGGTGGACTGGCAAGAGATCAAAATCTTCAA	540
Db	685	CTCTACTCTCTGAGAGGTTTTCGAAGATGTGGTGGACTGGCAAGAGATCAAAATCTTCAA	744
Qy	541	TCAAGAGACGGATCATTTTAAAGCTCCCTGCACTACTCTGTTGTCTTCAATGCACT	600
Db	745	TCAAGAGACGGATCATTTTAAAGCTCCCTGCACTACTCTGTTGTCTTCAATGCACT	804
Qy	601	CAAAAACAAACGATGCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGGGACTACGTT	660
Db	805	CAAAAACAAACGATGCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGGGACTACGTT	864
Qy	661	CTTGGCCATTAACCACTTGTATTTTGAAGCTCTGAGCGCTGCGATACAGTTGAAAGC	720
Db	865	CTTGGCCATTAACCACTTGTATTTTGAAGCTCTGAGCGCTGCGATACAGTTGAAAGC	924
Qy	721	TTGGGAATCGATCGCTATTTTCAAGAAAGAAATCAAAGAAATCTCTGGATTACGTTTATAGG	780
Db	925	TTGGGAATCGATCGCTATTTTCAAGAAAGAAATCAAAGAAATCTCTGGATTACGTTTATAGG	984
Qy	781	TACTGGGACGCGGAAAGAGCGTGGGATGGGCAAGATGCATCTCTTCTGATGTCGAT	840
Db	985	TACTGGGACGCGGAAAGAGCGTGGGATGGGCAAGATGCATCTCTTCTGATGTCGAT	1044
Qy	841	GACACTGCCATGGCTCTTAGAATCTCTGAGACTTCAATGATACAAATGTATCTTCAGATGT	900
Db	1045	GACACTGCCATGGCTCTTAGAATCTCTGAGACTTCAATGATACAAATGTATCTTCAGATGT	1104
Qy	901	CTGAGGAATTTTCAAGAGACGAGAAAGAGACTTCTTTTCTGCTTTCGCGGTCAACGCAATTT	960
Db	1105	CTGAGGAATTTTCAAGAGACGAGAAAGAGACTTCTTTTCTGCTTTCGCGGTCAACGCAATTT	1164
Qy	961	GGTGTGACCGATATCTTAACTTTATAGATGTTTCAAAAGTATGTTTTCGCGGAGAAAG	1020
Db	1165	GGTGTGACCGATATCTTAACTTTATAGATGTTTCAAAAGTATGTTTTCGCGGAGAAAG	1224
Qy	1021	ATAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGCCTCAAAAC	1080
Db	1225	ATAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGCCTCAAAAC	1284
Qy	1081	AACGCAATTTGATAGTGGGCTGTCAAGAGGATCTTCTGGAGAGGTGAGTATGCTATA	1140
Db	1285	AACGCAATTTGATAGTGGGCTGTCAAGAGGATCTTCTGGAGAGGTGAGTATGCTATA	1344
Qy	1141	AGATATCCGTGCAATAGATGATGCAAGATTTGGAGGCAAGAGTTTACATAGAGCAATTT	1200
Db	1345	AGATATCCGTGCAATAGATGATGCAAGATTTGGAGGCAAGAGTTTACATAGAGCAATTT	1404
Qy	1201	GGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAA	1260
Db	1405	GGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAA	1464
Qy	1261	AAATATTTGGAGCTGCGCAAAATTTGAGCTTCAATATGTTGAGGCGCTTACACCAAAAGGAG	1320
Db	1465	AAATATTTGGAGCTGCGCAAAATTTGAGCTTCAATATGTTGAGGCGCTTACACCAAAAGGAG	1524
Qy	1321	ACTCAACACATTTCTGAGCTGGTGGAGAAATCGGGATTCATGATCTTACATTCACCCGC	1380
Db	1525	ACTCAACACATTTCTGAGCTGGTGGAGAAATCGGGATTCATGATCTTACATTCACCCGC	1584
Qy	1381	CAGCGCGCTGTGAAATGATTTCTCAGTGGCGGTTAGTATGTTTGTAGCCAGAAATTCGCT	1440

Db 1585 CAGCGCCTGTGGAATGATATTCTCAGTGGCGGTAGTATGTTTGAGCCAGAAATTCGCT 1644  
QY 1441 GCTTGTAGAAATGCCCTATGCCAAGACTTCTTGCTCGCAGTTATCTAGACGATCTTTAC 1500  
Db 1645 GCTTGTAGAAATGCCCTATGCCAAGACTTCTTGCTCGCAGTTATCTAGACGATCTTTAC 1704  
QY 1501 GACACCCAGGATCTCTGGATGATCTTAAATGTTCTCTGAGCGGTCCGAGATGGAT 1560  
Db 1705 GACACCCAGGATCTCTGGATGATCTTAAATGTTCTCTGAGCGGTCCGAGATGGAT 1764  
QY 1561 ATCTCTGTCTGATAGCGTTCCGGATATCACTGGAAGTTTGTCTTCTAGGCGGTGATC 1620  
Db 1765 ATCTCTGTCTGATAGCGTTCCGGATATCACTGGAAGTTTGTCTTCTAGGCGGTGATC 1824  
QY 1621 AACACAGTGAATGATTTGGAAAGATGGAATCAAGAAACAAGCCGCTGATGTGGGC 1680  
Db 1825 AACACAGTGAATGATTTGGAAAGATGGAATCAAGAAACAAGCCGCTGATGTGGGC 1884  
QY 1681 TATCTTCGAAAGATGATGGGAGGCTTGTCTCGCATCTATACCAAGAGCCGAATGGTCG 1740  
Db 1885 TATCTTCGAAAGATGATGGGAGGCTTGTCTCGCATCTATACCAAGAGCCGAATGGTCG 1944  
QY 1741 GCAGCAAAATGATGTGCGCAATTCACAGATATGTGAAATGCCAAAGTGTCCATAGCA 1800  
Db 1945 GCAGCAAAATGATGTGCGCAATTCACAGATATGTGAAATGCCAAAGTGTCCATAGCA 2004  
QY 1801 CTTGGCAGTGTCTACTAACTCAATCTTTTTCACCTGGGAATTAATCTTCTGATTAAT 1860  
Db 2005 CTTGGCAGTGTCTACTAACTCAATCTTTTTCACCTGGGAATTAATCTTCTGATTAAT 2064  
QY 1861 TTACAGCAATGATGACCTTCCGTCCAAATTTCTGCATCTTGTGCTTTGACTGGACGACTA 1920  
Db 2065 TTACAGCAATGATGACCTTCCGTCCAAATTTCTGCATCTTGTGCTTTGACTGGACGACTA 2124  
QY 1921 ATCAATGACACCAAGACTTACCAAGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
Db 2125 ATCAATGACACCAAGACTTACCAAGCCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 2184  
QY 1981 CAGTGTACATAGGGGAAATCCGGATGTGCACAGAGAAAGACTCTTAAGTCAATGTTTAT 2040  
Db 2185 CAGTGTACATAGGGGAAATCCGGATGTGCACAGAGAAAGACTCTTAAGTCAATGTTTAT 2244  
QY 2041 GGTATCATCGACACGACCTGAAGGAATTTGAATTTGGAGTTGGCCAAACCCAGCGCAAT 2100  
Db 2245 GGTATCATCGACACGACCTGAAGGAATTTGAATTTGGAGTTGGCCAAACCCAGCGCAAT 2304  
QY 2101 GCCCAATGTGTGAGAGACTGCTGTTCAACACTGCAAGACTGATGCAAGCTGTTTAT 2160  
Db 2305 GCCCAATGTGTGAGAGACTGCTGTTCAACACTGCAAGACTGATGCAAGCTGTTTAT 2364  
QY 2161 ATGTACAGATGGCTTTGGTATCTCTGACAAAGATGAAGACCATGTGACCGCACT 2220  
Db 2365 ATGTACAGATGGCTTTGGTATCTCTGACAAAGATGAAGACCATGTGACCGCACT 2424  
QY 2221 CTTTTCGATCTGTGGCGTAG 2241  
Db 2425 CTTTTCGATCTGTGGCGTAG 2445

RESULT 4  
US-10-041-007-32  
; Sequence 32, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G.  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthese  
; FILE REFERENCE: P02081US1  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 2622  
; TYPE: DNA  
; ORGANISM: Ginkgo biloba  
US-10-041-007-32

Query Match 100.0%; Score 2241; DB 14; Length 2622;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGTTTTCAGTCCATGGCGACGCGTGAACGAAATTCATCTGCAATATGATACAGCTTGGTG 60  
Db 382 ATGTTTTCAGTCCATGGCGACGCGTGAACGAAATTCATCTGCAATATGATACAGCTTGGTG 441  
QY 61 GCAAGAAATTCGCTCAATTTGACGCGCTCTGGTGCAACCCCAATTTTCCCAACGCTTCAATGG 120  
Db 442 GCAAGAAATTCGCTCAATTTGACGCGCTCTGGTGCAACCCCAATTTTCCCAACGCTTCAATGG 501  
QY 121 ATTCTGAAACAATCAACTGCCAGATGGCTGTGGGGTGAGAGTGCATTTTCTGGCGTAT 180  
Db 502 ATTCTGAAACAATCAACTGCCAGATGGCTGTGGGGTGAGAGTGCATTTTCTGGCGTAT 561  
QY 181 GACAGAGTTTAAACAATCTCGCTCGCTCCTCACTCTCAAAATATGGAATAAGGGCGAC 240  
Db 562 GACAGAGTTTAAACAATCTCGCTCGCTCCTCACTCTCAAAATATGGAATAAGGGCGAC 621  
QY 241 ATTCAAGTGCAGAAAGGGGTGAGTTTGTGAGAAAAACACATGGAAGAAATGAAGGACGAA 300  
Db 622 ATTCAAGTGCAGAAAGGGGTGAGTTTGTGAGAAAAACACATGGAAGAAATGAAGGACGAA 681  
QY 301 GGTGACAATCAACAGGCAAGTGGATTCGAGGTGCTGTTCTGCAATTTAGATGAAGCA 360  
Db 682 GGTGACAATCAACAGGCAAGTGGATTCGAGGTGCTGTTCTGCAATTTAGATGAAGCA 741  
QY 361 AAAAGCTTGGGATTTGATCTTCTTATCACTCCCTTCACTCCCAAAATCCCAAAAG 420  
Db 742 AAAAGCTTGGGATTTGATCTTCTTATCACTCCCTTCACTCCCAAAATCCCAAAAG 801  
QY 421 CGCCAGAAAAAGCTTCAAAAGATTTCCCTCAATTTCTTCAATCAATCAAGCGCGTTG 480  
Db 802 CGCCAGAAAAAGCTTCAAAAGATTTCCCTCAATTTCTTCAATCAATCAAGCGCGTTG 861  
QY 481 CTCTACTCTCTGGAGGGTTTGAAGATGTGGTGGACTGGCAAGAGATCAAAATCTTCAA 540  
Db 862 CTCTACTCTCTGGAGGGTTTGAAGATGTGGTGGACTGGCAAGAGATCAAAATCTTCAA 921  
QY 541 TCAAGAGACGGATCAATTTTAAAGCTCCCTGCAATCTACTGCTTGTGCTTCAAGCACAT 600  
Db 922 TCAAGAGACGGATCAATTTTAAAGCTCCCTGCAATCTACTGCTTGTGCTTCAAGCACAT 981  
QY 601 CAAAAACAACGATGCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTAGCTT 660  
Db 982 CAAAAACAACGATGCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTAGCTT 1041  
QY 661 CTTTGCATTAACCACTTGAATTTGAACGCTCTGGGCTGTGATGATCAAGTGAACGC 720  
Db 1042 CTTTGCATTAACCACTTGAATTTGAACGCTCTGGGCTGTGATGATCAAGTGAACGC 1101  
QY 721 TTGGGATCGATCGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATTAAGTATAGG 780  
Db 1102 TTGGGATCGATCGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATTAAGTATAGG 1161  
QY 781 TACTGGAGCGCCGAAAGAGCGGTGGGATGGGCAAGATGCAATCTTCTTCTGATGTCGAT 840  
Db 1162 TACTGGAGCGCCGAAAGAGCGGTGGGATGGGCAAGATGCAATCTTCTTCTGATGTCGAT 1221  
QY 841 GACACTGCCATGGGTCTTAGAATCTTGAGATCTTCAAGTGAATCAATGATATCTTCAAGTGT 900  
Db 1222 GACACTGCCATGGGTCTTAGAATCTTGAGATCTTCAAGTGAATCAATGATATCTTCAAGTGT 1281  
QY 901 CTGGAGAAATTCAGACACGAGAAAGGAGACTTCTTTTGTCTTTGCGGTCAAGCGCAATTT 960

1282 CTGGAGAAATTTAGAGACGAGAAAGAGACTCTTTTGTCTTGGCGGTCAAAACGCAAAAT 1341  
961 GGTGTGACCGGATTAATCTTAACCTTTATAGATGTTCAAGATATGTTTTCGGGAGAAAG 1020  
1342 GGTGTGACCGGATTAATCTTAACCTTTATAGATGTTCAAGATATGTTTTCGGGAGAAAG 1401  
1021 ATAAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGGCCAAAAAC 1080  
1402 ATAAATGGAAGAGCTAAGACCTTCACTACAAATCATCTCCAAATGCTCTTGGCCAAAAAC 1461  
1081 AAGCATTCTTATAGTGGGCTGTCAAGAGGATCTTCTCGAGAGGTGGAGTATGCTATA 1140  
1462 AAGCATTCTTATAGTGGGCTGTCAAGAGGATCTTCTCGAGAGGTGGAGTATGCTATA 1521  
1141 AAGTATCCGTGGCANTAGAAGATATGCCAAGATATGGAGGCAAGAGTTACATAGAGCAATTT 1200  
1522 AAGTATCCGTGGCANTAGAAGATATGCCAAGATATGGAGGCAAGAGTTACATAGAGCAATTT 1581  
1201 GGATCAAAATGATGCTGGCTGGGGAAGACATGCTGTATAGATGCTATATGTGAGCAACGAA 1260  
1582 GGATCAAAATGATGCTGGCTGGGGAAGACATGCTGTATAGATGCTATATGTGAGCAACGAA 1641  
1561 AAATATTTGAGCTGGCCAAATTTGGACTTCAATATGTTGCGAGGCTTTACACCAAAAGGAG 1320  
1642 AAATATTTGAGCTGGCCAAATTTGGACTTCAATATGTTGCGAGGCTTTACACCAAAAGGAG 1701  
1321 ACTCAACACATGTCAGCTGTGAGAGAGATCGGATTCATGATGCTTACATTCACCCGC 1380  
1702 ACTCAACACATGTCAGCTGTGAGAGAGATCGGATTCATGATGCTTACATTCACCCGC 1761  
1381 CAGCGGCTGTGGAAATGTAATTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCT 1440  
1762 CAGCGGCTGTGGAAATGTAATTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCT 1821  
1441 GCTTTAGAAATGCTATGCTCAAGACATTTCTGCTCGCAGATTTCTAGAGCATTTTAC 1500  
1822 GCTTTAGAAATGCTATGCTCAAGACATTTCTGCTCGCAGATTTCTAGAGCATTTTAC 1881  
1501 GACACCCAGGATCTCTGGATGATCTTAAATGTTCTCTGAAGCGGTCGGAAGTGGAT 1560  
1882 GACACCCAGGATCTCTGGATGATCTTAAATGTTCTCTGAAGCGGTCGGAAGTGGAT 1941  
1561 ATCTGTGCTGGATGAGGTTTGGGATTAATCAGTTTGAAGTTTGTCTCTAGGGCTGTAC 1620  
1942 ATCTGTGCTGGATGAGGTTTGGGATTAATCAGTTTGAAGTTTGTCTCTAGGGCTGTAC 2001  
1621 AACACAGTGAATGGAATTTGGGAAAGATGGAATCAAGGACAGGCGGTGATGCTGGGC 1680  
2002 AACACAGTGAATGGAATTTGGGAAAGATGGAATCAAGGACAGGCGGTGATGCTGGGC 2061  
1681 TATCTTCGAAAGATATGGAGGGCTTGTCTCGCATCGTATACAAAGAGAGCCGAATGGTGC 1740  
2062 TATCTTCGAAAGATATGGAGGGCTTGTCTCGCATCGTATACAAAGAGAGCCGAATGGTGC 2121  
1741 GCAGCAAAAGTATGTGCCGACATTTCAACGAATATGTGGAAATGTCGCAAGTGTCCATAGCA 1800  
2122 GCAGCAAAAGTATGTGCCGACATTTCAACGAATATGTGGAAATGTCGCAAGTGTCCATAGCA 2181  
1801 CTTGGCAGTGTGATTAACCTCAATCTTTTCACTGGAGAAATTAATCTTCTGATTAATTT 1860  
2182 CTTGGCAGTGTGATTAACCTCAATCTTTTCACTGGAGAAATTAATCTTCTGATTAATTT 2241  
1861 TTACAGCAAGTATGACCTTCGGTCCAAATTTCTGCATCTTGTGCTTTGATGAGCACTA 1920  
2242 TTACAGCAAGTATGACCTTCGGTCCAAATTTCTGCATCTTGTGCTTTGATGAGCACTA 2301  
1921 ATCAATGACCAAGACTTACAGGCGGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTA 1980  
2302 ATCAATGACCAAGACTTACAGGCGGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTA 2361  
1981 CAGTGTACATGAGGGAATTCGGAGTGCACAGAGGAAGAGCTCTTAAGTCAATGTTTAT 2040  
2362 CAGTGTACATGAGGGAATTCGGAGTGCACAGAGGAAGAGCTCTTAAGTCAATGTTTAT 2421

2041 GGTATCATCGAACACGCACTGAGGAATTAATTGGAGTTGGCCAAACCCAGCGAGCAAT 2100  
2422 GGTATCATCGAACACGCACTGAGGAATTAATTGGAGTTGGCCAAACCCAGCGAGCAAT 2481  
2101 GCCCAATTTGTGTGAGAGAGCTGCTGTTCAACACATGCAAGAGTGTAGTGCAGCTGTTTTAT 2160  
2482 GCCCAATTTGTGTGAGAGAGCTGCTGTTCAACACATGCAAGAGTGTAGTGCAGCTGTTTTAT 2541  
2161 ATCTACAGAGATCGCTTTTGGTATCTCTGACAAAGAGATGAAAGACCAATGTGACCGCAACT 2220  
2542 ATGTACAGAGATCGCTTTTGGTATCTCTGACAAAGAGATGAAAGACCAATGTGACCGCAACT 2601  
2221 CTTTTCGATCCTGTGGCGTAG 2241  
2602 CTTTTCGATCCTGTGGCGTAG 2622

RESULT 5  
US-10-041-018-397  
; Sequence 397, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080US1/10025547  
; CURRENT APPLICATION NUMBER: US/10/041.018  
; PRIORITY FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 397  
; LENGTH: 2705  
; TYPE: DNA  
; ORGANISM: Ginkgo biloba  
US-10-041-018-397

Query Match 100.0%; Score 2241; DB 12; Length 2705;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTTTCAGTCCATGCGGCGAGCGTGAAGCAATCCATCTGCATATGATACAGCTTGGGTG 60  
404 ATGTTTTCAGTCCATGCGGCGAGCGTGAAGCAATCCATCTGCATATGATACAGCTTGGGTG 463  
61 GCAAGAAATTCGTCATTAATTGACGCTCTGTGTGCAACCCCAATTTCCCAAAACGCTTCAATGG 120  
464 GCAAGAAATTCGTCATTAATTGACGCTCTGTGTGCAACCCCAATTTCCCAAAACGCTTCAATGG 523  
121 ATCTGAACAATCAACTGCCAGATGCTGTGGGTGAGAGTGCATTTTCTGGCGGTAT 180  
524 ATCTGAACAATCAACTGCCAGATGCTGTGGGTGAGAGTGCATTTTCTGGCGGTAT 583  
181 GACAGAGTTTAAACACACTCTCGCTTCCCTCTCACTCTCAAAATATGGAATAAGGGCGAC 240  
584 GACAGAGTTTAAACACACTCTCGCTTCCCTCTCACTCTCAAAATATGGAATAAGGGCGAC 643  
241 ATCTAAGTGCAGAAAGGGGTGAGTTTGTGAGAAACACATGGAAGAAATGGAAGGAGCA 300  
644 ATCTAAGTGCAGAAAGGGGTGAGTTTGTGAGAAACACATGGAAGAAATGGAAGGAGCA 703  
301 GCTGCAAAATCAGAGCCCAAGTGGATTGAGTTCGAGGTGCTGTTCTTCCCAATGTTAGATGAAGCA 360  
704 GCTGCAAAATCAGAGCCCAAGTGGATTGAGTTCGAGGTGCTGTTCTTCCCAATGTTAGATGAAGCA 763  
361 AAAAGCTTGGGATTTGATCTTCTTATACCTCCCTTTTATCTCCCAATCCACCAAG 420  
764 AAAAGCTTGGGATTTGATCTTCTTATACCTCCCTTTTATCTCCCAATCCACCAAG 823  
421 CGCAGAGAAAGCTTCAAAAGATTCCTTCAATGTTCTTCAATACCAATCAGACGGCGGTG 480



Db 824 CGCAGAAAGCTTCAAGATTTCCCTCAATGTTCTTATACCATCATAGCGGTTG 383  
 Qy 481 CTCTACTCTCTGGAGGGTTGCAAGATGTTGGTGAATGCGCAAGATCAAAATCTTTCAA 340  
 Db 884 CTCTACTCTCTGGAGGGTTGCAAGATGTTGGTGAATGCGCAAGATCAAAATCTTTCAA 943  
 Qy 541 TCAAGAGACGGATCATTTTAAAGTCCCTGCACTACTACTGTTGTTCTCATGCACT 500  
 Db 944 TCAAGAGACGGATCATTTTAAAGTCCCTGCACTACTACTGTTGTTCTCATGCACT 1003  
 Qy 601 CAAAAAAGAGCGCTCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTAGTT 560  
 Db 1004 CAAAAAAGAGCGCTCTCCACTTTCTCAACTTCTGCTCAGCAAAATTTGGCGACTAGTT 1063  
 Qy 661 CTTGGCAATTAACCACTTGATCTATTTGAACGCTCTGGGCTGTCGATACAGTTGAACGC 720  
 Db 1064 CTTGGCAATTAACCACTTGATCTATTTGAACGCTCTGGGCTGTCGATACAGTTGAACGC 1123  
 Qy 721 TTGGGAATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTAAGTTATAGG 780  
 Db 1124 TTGGGAATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTAAGTTATAGG 1183  
 Qy 781 TACTGGACGCGCAAGAGCGGTGGGATGGCAAGATGCAATCTATTTCTGATGTCGAT 340  
 Db 1184 TACTGGACGCGCAAGAGCGGTGGGATGGCAAGATGCAATCTATTTCTGATGTCGAT 1243  
 Qy 841 GACATCGCCATGGGTCTTAGAATCTCTAGATCTCTGAGATCAATGATCTTCAAGATGTT 900  
 Db 1244 GACATCGCCATGGGTCTTAGAATCTCTGAGATCTCTGAGATCAATGATCTTCAAGATGTT 1303  
 Qy 901 CTGGAGAAATTTCAAGACGAGAAAGGAGACTTCTTTTGGCTTTGCGGCTCAAGCGCAATTT 360  
 Db 1304 CTGGAGAAATTTCAAGACGAGAAAGGAGACTTCTTTTGGCTTTGCGGCTCAAGCGCAATTT 1363  
 Qy 961 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1020  
 Db 1364 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1423  
 Qy 1021 ATATGGAAGAGCTAAGACTTCACTCAAAATCATCTCAAAATGCTCTTGGCAAAAC 1080  
 Db 1424 ATATGGAAGAGCTAAGACTTCACTCAAAATCATCTCAAAATGCTCTTGGCAAAAC 1483  
 Qy 1081 AAGCATTTGATAGTGGCTGTCAAGAGATCTTCTGGAGAGGTGGATATGCTATA 1140  
 Db 1484 AAGCATTTGATAGTGGCTGTCAAGAGATCTTCTGGAGAGGTGGATATGCTATA 1543  
 Qy 1141 AAGTATCCGTGGCATAGAGTATGCCAAGATTTGAGGCAAGAGTTACATAGACAAATTT 1200  
 Db 1544 AAGTATCCGTGGCATAGAGTATGCCAAGATTTGAGGCAAGAGTTACATAGACAAATTT 1603  
 Qy 1201 GGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAA 1260  
 Db 1604 GGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAA 1663  
 Qy 1261 AAATATTTGAGCTGGCCAAATTTGGAATTCATATGTTGAGGCTTACACCAAAAGGAG 1320  
 Db 1664 AAATATTTGAGCTGGCCAAATTTGGAATTCATATGTTGAGGCTTACACCAAAAGGAG 1723  
 Qy 1321 ACTCAACACATTTGCTGGTGGAGAGATCGGATTCATGATCTTACATTCACCCCGC 1380  
 Db 1724 ACTCAACACATTTGCTGGTGGAGAGATCGGATTCATGATCTTACATTCACCCCGC 1783  
 Qy 1381 CAGCGGCTGTGGAATATGTTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGATTCGCT 1440  
 Db 1784 CAGCGGCTGTGGAATATGTTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGATTCGCT 1843  
 Qy 1441 GCTTGTAGATGCTGATGCAAGACTTCTTGGCTCTGCGAGTTATTTAGACGATCTTTTAC 1500  
 Db 1844 GCTTGTAGATGCTGATGCAAGACTTCTTGGCTCTGCGAGTTATTTAGACGATCTTTTAC 1903  
 Qy 1501 GACACCCACGGATCTCTGGATGATCTTAATTTGTTCTCTGAAGCGGTCCGAAGATGGAT 1560  
 Db 1904 GACACCCACGGATCTCTGGATGATCTTAATTTGTTCTCTGAAGCGGTCCGAAGATGGAT 1963

Qy 1561 ATCTCTGTCTGTAGTAGCGTTTCGGGATTAATCAGTTGAAAGTTTGCTTCTTAGGCTGTAC 1620  
 Db 1964 ATCTCTGTCTGTAGTAGCGTTTCGGGATTAATCAGTTGAAAGTTTGCTTCTTAGGCTGTAC 2023  
 Qy 1621 AACACAGTGAATCGATTTTGGAAAGATGGAATCAAGGACCAAGCGCGTGAATGTCTGGGC 1680  
 Db 2024 AACACAGTGAATCGATTTTGGAAAGATGGAATCAAGGACCAAGCGCGTGAATGTCTGGGC 2083  
 Qy 1681 TATCTTCGAAAGATATGGGAGGCTTCTGTCGATTCGTATACCAAGAACGCGAATGTCTG 1740  
 Db 2084 TATCTTCGAAAGATATGGGAGGCTTCTGTCGATTCGTATACCAAGAACGCGAATGTCTG 2143  
 Qy 1741 GCAGCAAGATATGTCGCAATTCACGATATATGCGAAATATGCGAAATATGCGAAATATG 1800  
 Db 2144 GCAGCAAGATATGTCGCAATTCACGATATATGCGAAATATGCGAAATATGCGAAATATG 2203  
 Qy 1801 CTTGGCAGAGTCTGATCTAAACTCAATCTTTTCACTGGAGAAATTAATCTCTGATTAAT 1860  
 Db 2204 CTTGGCAGAGTCTGATCTAAACTCAATCTTTTCACTGGAGAAATTAATCTCTGATTAAT 2263  
 Qy 1861 TTAACAGCAAGTAGACCTTCCGTCCTCAAAATTTCTGATCTTTGACTTGGACGACTA 1920  
 Db 2264 TTAACAGCAAGTAGACCTTCCGTCCTCAAAATTTCTGATCTTTGACTTGGACGACTA 2323  
 Qy 1921 ATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTCCAGCGTA 1980  
 Db 2324 ATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTCCAGCGTA 2383  
 Qy 1981 CAGTGTACATGAGGGAATTCGGAGTGCACAGAGAGAGAAAGCTCTAAGTCATGTTTAT 2040  
 Db 2384 CAGTGTACATGAGGGAATTCGGAGTGCACAGAGAGAGAAAGCTCTAAGTCATGTTTAT 2443  
 Qy 2041 GGTATCATCGACCAACGACTTGAAGGAAATTTGAAATTTGGAGTTGGCCAAACCCAGCGCAAT 2100  
 Db 2444 GGTATCATCGACCAACGACTTGAAGGAAATTTGAAATTTGGAGTTGGCCAAACCCAGCGCAAT 2503  
 Qy 2101 GCGCCATTTGTGTGAGAGACTGCTGTTCACACATGCAAGAGTGAATGCAAGCTGTTTAT 2160  
 Db 2504 GCGCCATTTGTGTGAGAGACTGCTGTTCACACATGCAAGAGTGAATGCAAGCTGTTTAT 2563  
 Qy 2161 ATGTACAGAGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2220  
 Db 2564 ATGTACAGAGATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2623  
 Qy 2221 CTTTTCGATCTCTGGCGTAG 2241  
 Db 2624 CTTTTCGATCTCTGGCGTAG 2644

RESULT 6

US-10-041-007-1  
 ; Sequence 1, Application US/10041007  
 ; Publication No. US20020164736A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matsuda, Seiichi P.T.  
 ; APPLICANT: Schepmann, Hala G.  
 ; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthese  
 ; FILE REFERENCE: P02081US1  
 ; CURRENT APPLICATION NUMBER: US/10/041,007  
 ; PRIOR FILING DATE: 2002-01-07  
 ; PRIOR APPLICATION NUMBER: US 60/259,881  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: Patentin version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 2705  
 ; TYPE: DNA  
 ; ORGANISM: ginkgo biloba  
 ; US-10-041-007-1

Query Match 100.0%; Score 2241; DB 14; Length 2705;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 2241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	ATGTTTCAGTCCATGGCGGCGGTAAGCAATCCATCTGCATATCATATCAGCTTCGGGTG 60
Db	404	ATGTTTCAGTCCATGGCGGCGGTAAGCAATCCATCTGCATATCATATCAGCTTCGGGTG 463
QY	61	GCAAGAAATCCGTCATTTGACGGCTCTGCTGTCACCCCAATTTCCCAACGCTTCAATGG 120
Db	464	GCAAGAAATCCGTCATTTGACGGCTCTGCTGTCACCCCAATTTCCCAACGCTTCAATGG 523
QY	121	ATTCTGAACAATCAATCGCCAGATGGCTCGTGGGGTGGAGGATGCAATTTTCTGGGGTAT 180
Db	524	ATTCTGAACAATCAATCGCCAGATGGCTCGTGGGGTGGAGGATGCAATTTTCTGGGGTAT 583
QY	181	GACAGAGTTTAAACACTCTCGCTGCTCCTCTCACTCTCAAAATATGGAATATAGGGCGAC 240
Db	584	GACAGAGTTTAAACACTCTCGCTGCTCCTCTCACTCTCAAAATATGGAATATAGGGCGAC 643
QY	241	ATTCAGTGCAGAAAGGGGTGAGTTTGTGAGAAACACATGGAAGAAATGAAGGAGAA 300
Db	644	ATTCAGTGCAGAAAGGGGTGAGTTTGTGAGAAACACATGGAAGAAATGAAGGAGAA 703
QY	301	GCTGACAAATCACAGGCCAAGTGGATTCGAGTCTGTTCTGCAATGTTAGATGAAGCA 360
Db	704	GCTGACAAATCACAGGCCAAGTGGATTCGAGTCTGTTCTGCAATGTTAGATGAAGCA 763
QY	361	AAAAGCTTGGGATTTGGATCTTCTTATCATCCTCCCTTTCATCTCCCAATCCACAAAG 420
Db	764	AAAAGCTTGGGATTTGGATCTTCTTATCATCCTCCCTTTCATCTCCCAATCCACAAAG 823
QY	421	CGCAGAAAGAGCTTCAAAAGATTCCTCAATGTTCTTCAATACCATCAGACGGCGTTG 480
Db	824	CGCAGAAAGAGCTTCAAAAGATTCCTCAATGTTCTTCAATACCATCAGACGGCGTTG 883
QY	481	CTCTACTCTCTGGAGGTTTGCAGATGTTGGTGAATGCTGCAAGAGATCAAAATCTTCAA 540
Db	884	CTCTACTCTCTGGAGGTTTGCAGATGTTGGTGAATGCTGCAAGAGATCAAAATCTTCAA 543
QY	541	TCAAGAGACGATCAATTTTAAGCTCCCTGATCTACTGCTGCTCAGCAAAATTTGGGACTACGTT 600
Db	944	TCAAGAGACGATCAATTTTAAGCTCCCTGATCTACTGCTGCTCAGCAAAATTTGGGACTACGTT 1003
QY	601	CAAAACAAACGATGCTCCACTTTCTCAACTCTGCTGCTCAGCAAAATTTGGGACTACGTT 660
Db	1004	CAAAACAAACGATGCTCCACTTTCTCAACTCTGCTGCTCAGCAAAATTTGGGACTACGTT 1063
QY	661	CCTTGCATTTACCCATTCATCTAATTTGAACGCTCTGGGCTGTCATACAGTTGAACGC 720
Db	1064	CCTTGCATTTACCCATTCATCTAATTTGAACGCTCTGGGCTGTCATACAGTTGAACGC 1123
QY	721	TTGGGATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTCAGTTATAGG 780
Db	1124	TTGGGATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATTCAGTTATAGG 1183
QY	781	TACTGGGACGCCGAAAGAGCGTGGGATGGCAAGATGCAATCTTATCTGATGTCGAT 840
Db	1184	TACTGGGACGCCGAAAGAGCGTGGGATGGCAAGATGCAATCTTATCTGATGTCGAT 1243
QY	841	GACATGCCATGGGTCTTAGAATCTGAGATCTTCATGGATACAAATCTATCTTCCAGATGTT 900
Db	1244	GACATGCCATGGGTCTTAGAATCTGAGATCTTCATGGATACAAATCTATCTTCCAGATGTT 1303
QY	901	CTGGAGAAATTCAGAGACGAGAAAGAGACTCTTTTGTCTTTCGGGTCAACGCAAAAT 960
Db	1304	CTGGAGAAATTCAGAGACGAGAAAGAGACTCTTTTGTCTTTCGGGTCAACGCAAAAT 1363
QY	961	GGTGTGACCAATTAATCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1020
Db	1364	GGTGTGACCAATTAATCTTTATAGATGTTTCAAGATGTTTTCGGGAGAAAG 1423
QY	1021	ATATGGAAGAGCTTACATCAAAATCATCTCCTCAAAATGCTCTTGGCAGAAAC 1080
Db	1424	ATATGGAAGAGCTTACATCAAAATCATCTCCTCAAAATGCTCTTGGCAGAAAC 1483

QY	1081	AACGCAATTTGATAAGTGGCTCTCAAGAGGATCTTCTCGAGAGGTGAGTATGCTATA 1140
Db	1484	AACGCAATTTGATAAGTGGCTCTCAAGAGGATCTTCTCGAGAGGTGAGTATGCTATA 1543
QY	1141	AAGTATCCGTGGCATAGAAATGATCCAGATTTGGAGGCAAGAAATTTACATAGAGCAATTT 1200
Db	1544	AAGTATCCGTGGCATAGAAATGATCCAGATTTGGAGGCAAGAAATTTACATAGAGCAATTT 1603
QY	1201	GGATCAAAATGATGTCCTGGCTGGGAGAGCTGTGTATPAAGATGCTATATGTGACCAACGAA 1260
Db	1604	GGATCAAAATGATGTCCTGGCTGGGAGAGCTGTGTATPAAGATGCTATATGTGACCAACGAA 1663
QY	1261	AAATATTTGGAGCTGGCCCAAAATTTGGACTTCAATATGCTGAGGCTTTACACCAAAAGGAG 1320
Db	1664	AAATATTTGGAGCTGGCCCAAAATTTGGACTTCAATATGCTGAGGCTTTACACCAAAAGGAG 1723
QY	1321	ACTCAACACATTTGCTAGCTGGTGGAGAGATCGGATTCATATGATCTTACATTCACCCGC 1380
Db	1724	ACTCAACACATTTGCTAGCTGGTGGAGAGATCGGATTCATATGATCTTACATTCACCCGC 1783
QY	1381	CAGGGCTCTGGAATGTAATTTCTCAGTGGCGGTTAGTATGTTTGGAGCAGAAATTCGCT 1440
Db	1784	CAGGGCTCTGGAATGTAATTTCTCAGTGGCGGTTAGTATGTTTGGAGCAGAAATTCGCT 1843
QY	1441	GCTTGTAGAAATTCCTATGCAAGACATTTCTGCTGCGAGTATTTCTAGACGATCTTTAC 1500
Db	1844	GCTTGTAGAAATTCCTATGCAAGACATTTCTGCTGCGAGTATTTCTAGACGATCTTTAC 1903
QY	1501	GACACCAACGATCTGGAATGATCTTAAATTTGTTCTGAAAGCGTCCGAAAGATGGAT 1560
Db	1904	GACACCAACGATCTGGAATGATCTTAAATTTGTTCTGAAAGCGTCCGAAAGATGGAT 1963
QY	1561	ATCTCTGTGCTGGATAGCTTCCGGATTAATCAGTTGAAAGTTTCTCTTACGGCTGTAC 1620
Db	1964	ATCTCTGTGCTGGATAGCTTCCGGATTAATCAGTTGAAAGTTTCTCTTACGGCTGTAC 2023
QY	1621	AACACAGTGAATGATTTGGAAAGATGGAATCAAGGAAACGAGCGGTGATGTCGGGC 1680
Db	2024	AACACAGTGAATGATTTGGAAAGATGGAATCAAGGAAACGAGCGGTGATGTCGGGC 2083
QY	1681	TATCTTCCGAAAGTATGGAGGGCTTGTCTGCAATCGTATACCAAGAAAGCGGATGTCG 1740
Db	2084	TATCTTCCGAAAGTATGGAGGGCTTGTCTGCAATCGTATACCAAGAAAGCGGATGTCG 2143
QY	1741	GCAGCAAAATGATGTCGCCACATTTCAACGAATATGTTGAAATGCCAAAGTGTCCATAGCA 1800
Db	2144	GCAGCAAAATGATGTCGCCACATTTCAACGAATATGTTGAAATGCCAAAGTGTCCATAGCA 2203
QY	1801	CTTGGCAGACGTCGTAATAAATCTCAATCTTTTCTTCTGAGAGATTAATCTTCTGATTAAT 1860
Db	2204	CTTGGCAGACGTCGTAATAAATCTCAATCTTTTCTTCTGAGAGATTAATCTTCTGATTAAT 2263
QY	1861	TTACAGCAAGTATAGACCTTGGTCCAAATTTCTGCAATCTTGTCTTGTGATGAGCACTA 1920
Db	2264	TTACAGCAAGTATAGACCTTGGTCCAAATTTCTGCAATCTTGTCTTGTGATGAGCACTA 2323
QY	1921	ATCAATGACACCAAGATTTACCAAGCGGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTA 1980
Db	2324	ATCAATGACACCAAGATTTACCAAGCGGAGAGAAACCGTGGTGAATTTGTTTCCAGCGTA 2383
QY	1981	CAGTGTACATGAGGAAATCCGGAGTGCACAGAGGAAAGAGCTCTAAGTCAATGTTAT 2040
Db	2384	CAGTGTACATGAGGAAATCCGGAGTGCACAGAGGAAAGAGCTCTAAGTCAATGTTAT 2443
QY	2041	GGTATCATCGACCAACGCTGAGGAAATTTGAAATTTGGAGTTGGCCCAACCCAGCGAGCAAT 2100
Db	2444	GGTATCATCGACCAACGCTGAGGAAATTTGAAATTTGGAGTTGGCCCAACCCAGCGAGCAAT 2503
QY	2101	GCCCAATTTGTGTGAGAGAGCTGCTGTTTCAACACACCTGCAAGAGTGTGAGCTGTTTAT 2160
Db	2504	GCCCAATTTGTGTGAGAGAGCTGCTGTTTCAACACACCTGCAAGAGTGTGAGCTGTTTAT 2563

QY 2161 ATGTACAGAGATGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTCCAGCGAAT 2220  
Db 2564 ATGTACAGAGATGCTTTGGTATCTCTGACAAAGAGATGAAGACCATGTCCAGCGAAT 2623  
QY 2221 CTTTTCGATCTGTGGCGTAG 2241  
Db 2624 CTTTTCGATCTGTGGCGTAG 2644

## RESULT 7

US-09-887-586A-55  
; Sequence 55, Application US/09887586A  
; Patent No. US20020094556A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: NO. US20020094556A1, Joseph P.  
; APPLICANT: Starke, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/887,586A  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(2606)  
; OTHER INFORMATION: abietadiene synthase  
US-09-887-586A-55

Query Match 48.1%; Score 1078; DB 9; Length 2861;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

QY 1 ATGTTTCAGTCCATGGCGGAGCGGTGAACCGAATCCATCTGCAATATGATACAGCTTGGGTG 50  
Db 378 ATGTTTAGATGTATGGGCTATGGCGAAGCGAATCCCTCTGCAATATGACACTGCTTGGGTA 437  
QY 61 GCAAGAAATCCCGTCAATTTGACGGCTCTGGTGCACCCCAATTTCCCAAGCGTTCATATGG 120  
Db 438 GCAAGAAATCCAGAGTTGATGGCTCTGCAACCCCTCACTTCTCTGAGACGGTTGAATGG 497  
QY 121 ATTCTGAAACAATCAATGCGAGATGGCTCTGGGGGTGAGGAGTGCATTTTCTGGCGGTAT 180  
Db 498 ATTCTCAAAATCAGTTGAAGATGGCTCTGGGGGTGAGGAGTCTACTTCTTGGCATAT 557  
QY 181 GACAGAGTTTAAACAATCTGCGCTCTCACTCTCAAAATATGGAATTAAGGCGAC 240  
Db 558 GACAGAAATCTGGCTACATCTGCATATGATATTATTAACCCCTTACCCTCTGGCGGTACTGGGGAG 517  
QY 241 ATTCAAGTCAGAAAGGGTGTGATTTGTGAGAAACACATGGAAGAAATGAAGACGAA 300  
Db 618 ACAAGATGACAGAAAGGTATTGAATTTCTCAGACACAAAGCTGGAAGATGGAAGATGAA 577  
QY 301 GCTGACAATCAGACGGCCAGTGTGATTCGAGGTCTGTTTCTGCAATGTTAGATGAAGCA 360  
Db 678 GCTGATAGTCATAGGCCAAGTGAATTTGAATAGTATTTCTGCAATGCTTAAGGAGCT 737  
QY 361 AAAAGCTTGGGATTTGATCTTCCTTATCACCTCCCTTTCATCTCCCAATCCCAAAAG 420  
Db 738 AAAATCTTAGGCTTGGATCTGCTTACGATTTTGCAATTCCTGCAACCAATCATCGAAAG 797  
QY 421 CGCCAGAAAAGCTTCAAAAGATTCCTCCCTCAATGTTCTTCAATCAACCATCAGACGGCGTTG 480

Db 798 CGGAGGCTAAGCTTAAAGGATTCACCACTGATGTTCTCTATGCCCTTCCAAACAGCTTA 857  
QY 481 CTCTACTCTCTGGAGGGTTTGCAGAGATGTGGTGGAGTGGCAAGAGATCACAATAATCTTCAA 540  
Db 858 TTGTATTTCTTGGAGGTTTACAAAGAAATAGTAGACTGGCAGAAATAATGAAACTTCAA 917  
QY 541 TCAAGAGACGGATCATTTTTAAGCTCCCTGCGATCTACTGCTGTGTGCTTCTCATGCACACT 600  
Db 918 TCCAAGGATGGATCATTTCTCAGCTCTCCGGCATCTACAGCGGCTGTATTTATCATGCTACA 977  
QY 601 CAAAACAAACGATGCTCCACTTTCTCAACTTCGTCCTCAGCAAAATTTGGCGACTACGTT 660  
Db 978 GGGAAACAAAGTGGCTTGGATTTCTTGAACCTTTGTCTTGAAGAAATTCGGAACCATGTG 1037  
QY 661 CTTTGCCATTTACCCATTTGATCTATTGAAACGCTCTGGGCTGTGCTATACAGTTTGAACGC 720  
Db 1038 CTTTGTCACTATCCGCTTGATCTATTGAAACGTTTGGCGGTTGATACAGTTTGAACGC 1097  
QY 721 TTGGGAATCGATCGCTATTTCAGAAAGAAATCAAGAAATCTCTGGATTTAGTTATAGG 780  
Db 1098 CTAGGTATCGATCGCTATTTCAGAAAGGAGATCAAGGAAGCATTTGGATTTATGTTTACAGC 1157  
QY 781 TACTGGGACGCCAAAGAGCGGTGGATGGCAAGATGCAATCTATTTCTGTATGTCGAT 840  
Db 1158 CATTTGGA---CGAAAGAGCATTTGGATGGCGAGAGAGAAATCTGTCTCTGTATTTGAT 1214  
QY 841 GACACTGCCATGGGTCTTAGAATCTCTGAGACTTCTGAGATACAAATGTATCTTCAGATGTT 900  
Db 1215 GATACAGCATGGGCTTCGAATCTTGAAGATTAATGATGATACATCTCTCAGATGTT 1274  
QY 901 CTGGAGATTTCAAGACGAGAAAGAGACTCTCTTTTCTGCTTGGCGGTCAAAACGCAATTT 960  
Db 1275 TTAATAACATTTAGAGATGAGAAATGGGAGTCTCTTTTGTCTTCTTGGGTCAAAACAGAGA 1334  
QY 961 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGTATGTTTTCGGGAGAAAG 1020  
Db 1335 GGAGTTACAGACATGTTTAAACGTTCAATCGTTGTTCACATGTTTCAATTTCCGGGAGAAACG 1394  
QY 1021 ATAAATGGAAGAGCTAAGACCTTCTCACTACAAATCATCTCCAAATGCTCTTGGCCAAAC 1080  
Db 1395 ATCATGGAAGAGCAAAACCTCTGTCGAAAGGTATCTGAGGAATGCTCTGGAATAATGTG 1454  
QY 1081 AAGCAATTTGATTAAGTGGGCTGTCAAGAGGATCTTCTCGGAGAGGTGAGATGCTATATA 1140  
Db 1455 GATGCTCTTTGACAAATGGGCTTTTAAAGAAATTTCCGGGAGAGGTAGATGCTACTC 1514  
QY 1141 AAGTATCCGTGGCATAGAAGTATGCCAAGATTCGAGGCAAGAGTTACATAGAGCAATTT 1200  
Db 1515 AATATCCCTGGCATAGAGATGTCACAGGTTGGAGGCTAGAAGCTATATTGAAACATAT 1574  
QY 1201 GGATCAAAATGATGCTCTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTGAGCAACGAA 1260  
Db 1575 GGCCAGATGATGTTGGCTTGGAAACCTGTATATATGATGCCATACATTTCCGAATGAA 1634  
QY 1261 AATATTTGGAGCTGGCCAAATTTGGAATTTCAATATGTTGGCAGGCTTTACACCAAGAGAG 1320  
Db 1635 AAGTATTTAGAACTAGCGAACTGAGCTTCAATTAAGGTGCACTTATACACCAACAGAG 1694  
QY 1321 ACTCAACACATTTGCTGAGCTGGGAGAGATCGGATTCGAATCTTACATCTTACATTCACCCGC 1380  
Db 1695 CTTCAAGATCTTGAAGGTGGTGAATCATCCGGTTTACGGATCTGAAATTTCACTCGT 1754  
QY 1381 CAGCGGCTGTGGAATGTTATTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCT 1440  
Db 1755 GAGCGTGTGACGGAATATATTTCTCACCGCATCTTTTATCTTTGAGCCCGAGTTTCT 1814  
QY 1441 GCTTGTAGAAATTTGCTTATGCAAGACTTCTTGTGCTCGAGTATTTCTAGACGATCTTTTAC 1500  
Db 1815 AAGTGACAGAGAGGTTTATACAAAACCTTCCAAATTTCACTGTTATTTTAGATGATCTTTAT 1874  
QY 1501 GACACCAACGATCTCTGGATGATCTTAAATGTTTCTCTGAAAGGCTCCGAAGTGGAT 1560

1875 GACGCCATGATCTTTAGACGATCTTAAGTTGTTACAGAAATCAGTCAAAAGATGGAT 1934  
1561 ATCTGTGCTGGATAGCTGCGGATAATCAGTTGAAAGTTTCTTCTAGGCTGTAC 1620  
1935 CTATCACTAGTGA-----CNAATGCCACAAATGAAATATGTTTGTGGTTTCTAC 1991  
1621 AACACAGTGAATGGATTTGGAAAAGATGACCTCAAGAAACAGGCCGCTGATGCTGGGC 1680  
1992 NATACTTTTAAATGATATAGCMAAAGAGGACGCTGAGAGGCAAGGGCGGATGTGCTAGGC 2051  
1681 TATCTTCGAAAAGTATGGAGGCTTGTCTCGCATCTATATACAAAGAACCGCAATGGTCG 1740  
2052 TACATTCAAAATGTTTGGAAAAGTCCAACTTGAAGCTTTACAGAAAGAGCAGAAATGGTCT 2111  
1741 GCAGCAAGATGATGTCGCGCATTTCAACGAATATGTGGAAAATCCAAAGTGTCCATAGCA 1800  
2112 GAAGCTAAATATGTGCCATCTTCAATGATATACATAGAGAATCGAGTGTCTCAATAGCA 2171  
1801 CTTGCGCAGCTGATCTAAACTCAATCTTTTCACTGGAGAAATTAATCTTCTGATTACAT 1860  
2172 TTGGGAACAGTCTGTTCTCAATAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA 2231  
1861 TTACAGCAAGTATGACCTTCCGCTCAGATTTCTGCTCTTGTCTTTGACTGGACGACTA 1920  
2232 CTCTCCAAAATTTGATCGGAACTTAGAATTTCTTCACTCATGGCTTAAACAGGGCGTTG 2291  
1921 ATCAATGACACCAAGACTTTACAGCGCGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
2292 GTGAATGACACCAAACTTTATCAGCGAGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATA 2351  
1981 CAGTGTATCATGAGGAAATCCGGAGTGACAGAGGAGAGCTCTAAAGTCAATGTTTAT 2040  
2352 CAATGTTTATATGAAGGACCATCTTAAATCTCTGAAGAAGAGCTCTTACACATGTCTAT 2411  
2041 GGTATCATCGACAACGCACTCAAGGAATTTGAATTTGGGAGTTGGCCAAACCCAGCAGCAAT 2100  
2412 AGTGTATGGAATATGCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAAA 2465  
2101 GCCCATTTGTGTGAGAAAGTCTGTTTCAACACTGCAAGAGTGCAGCTGTTTAT 2160  
2466 ATACCGGATATTTACAAAGACTGGTTTTTGAATCTGCAAGAAATTAATGCAACTCTTTAT 2525  
2161 ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCATGTCCAGCGGA 2217  
2526 ATGCAAGGGGATGGTTTGAACATATACATGATATGGAATTAAGAGCATGTCAAAAAT 2585  
2218 ACTCTTTTCGATCTCTGTGGCGTAG 2241  
2586 TGCCTCTTCCACCAACAGTTGCCTAG 2609

RESULT 8

US-09-903-012-55  
; Sequence 55, Application US/09903012  
; Patent No. US20020094557A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20020094557A1, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/903, 012  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58

FASTA: FaetsRQ for Windows Version 3.0

; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(2606)  
; OTHER INFORMATION: abietadiene synthase  
; US-09-903-012-55

Query Match 48.1%; Score 1078; DB 9; Length 2861;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

1 ATGTTTTCAGTCCATGGCGACGCTGAAAGCAATCCATCTGCATATGATACAGCTTGGGTG 60  
378 ATGTTTAGATGATGGCTATGGGAAACGAATCCCTCTGTCATATGACATGCTGCTGGGTA 437  
61 GCAAGAAATTCGCTCAATGACGGCTCTGTGTCACCCCAATTTCCCAAGGCTTCAATGG 120  
438 GCAAGGATTCACGACGCTTGTATGGCTCTGACCAACCTTCACCTTCTGAGACGGTTGAATGG 497  
121 ATTCTGAACAACTCAACTGCCAGATGGCTGCTGGGCTGAGGAGTGCAATTTTCTGGCGTAT 180  
498 ATTCTTCAAAATCAGTTGAAAGATGGCTCTGGGCTGAGGAGTCTTACCTTCTGGCATAT 557  
181 GACAGAGTTTAAACACTCTCGCTCTCCTCTCACTCTCAAAATATGGAATTAAGGGCGAC 240  
558 GACAGATATCTGGCTACACTTGCATGTATTTATACCTTACCTCTGCGGTACTGGGGAG 617  
241 ATTCAAGTCAGAAAGGGTTGATTTGTGAGAAACACATGGAAGAAATGAAGGACAA 300  
618 ACACAGTCAGAAAGGATTTGAATTTCTCAGGACACACAGCTGGAAGATGGAAGTGA 677  
301 GCTGACATCACAGGCCAAAGTGGATTCGAGTCTGTTTCTGCAATGTTAGATGAAGCA 360  
678 GCTGATAGTCAAGGCCAGTGGATTTGAAATAGTATTTCTGCAATGCTAAAGGAAGCT 737  
361 AAAAGCTTGGGATTTGATTTCTTATCATCCTCTCTTTCATCTCCCAATCCACCAAG 420  
738 AAAATCTTAGGCTTGGATCTGCTTACGATTTTGGCATTTCTTGAAGAAATTCGAAACCA 797  
421 CGCAGAAAGAGCTTCAAAAGATTTCCCTCAATGTTCTTCAATGCTTCAACCATCAGACGG 480  
798 CGGAGGCTAAGCTTTAAAGGATTTCCCACTGATGTTCTCTATGCCCTTCCAACACGTTA 857  
481 CTCTACTCTCTGGAGGGTTTGCAGATGTGTGAGTCTGGAGTCTGGCAAGAGATCACAAATCTTCAA 540  
858 TTGTATTTCTTTGGAAGGTTTACAGAAATAGTAGACTGGCAGAAATTAATGAACCTTCAA 917  
541 TCAAGAGACGGATCATTTTAAAGCTCCCTGCTCATCTACTGTTGTGTTCTTCTCATGCACACT 600  
918 TCCAGGATGATCATTTTCTCAGCTCTCCGGCATCTACAGCGGCTGTATTTCTATCGGTACA 977  
601 CAAACAAACGATGCTCCACTTTCTCAACTTTCTGCTCAGCAAAATTTGGCGACTACGTT 660  
978 GGGAAACAAAGTGTGTTGATTTCTTGAACCTTCTGAGTCTGCTGATACAGTGAACGC 720  
661 CTTTGGCAATACCACTTGTATTTGAACGCTCTGGGCTGCTGATACAGTGAACGC 720  
1038 CTTTGTCACTATCCGCTTGTATTTGAACGTTTGTGGGCGGTTGATACAGTGAACGC 1097  
721 TTGGGATCGATCGCTTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATACGTTTATAGG 780  
1098 CTAGGTATCGATCGCTTATTTCAAGAGGAGATCAAGGAGCATTTGGATTTATGTTACAGC 1157  
781 TACTGGGACGCCGAAAGAGCGGTGGGATGGGCAAGATGCAATCTTCTTCTGATGTCAT 840  
1158 CATTTGGA---CGAAAGAGGCAATGATGGGCGAGAGAGATCTGTCTCTGATTTGAT 1214  
841 GACATGCCATGGGTCTTAGATCTTCAATCTTCAAGTATGATGATGATCTTCTCAGATGTT 900

Db 1215 GATACAGCATGGGCTTCGAATCTTGAGATTACATGATACATGATCTCAGATGTT 1274  
Qy 901 CTGGAGAAATTCAGAGACGAGAAAGAGACTCTCTTTTGGCTTTCGGCTCAACCCAAAT 360  
Db 1275 TTAAGAAACATTTAGAGATGAGAAATGGGAGTCTCTTTTGGCTTTCGGCTCAACACAGAGA 1334  
Qy 961 GGTGTGACCGATATCTTAACCTTTATAGATGTTTCAAGATGATCTTTTCCGGAGAAAG 1020  
Db 1335 GGAGTTACAGACATGTTAAAGCTCAATCTGTTTTCATGTTTCAATCTCCGGAGAAACG 1394  
Qy 1021 ATATGGAAGAGCTAAGACCTTCCTCACTCAAAATCATCTCCAAATGCTCTTTCGCAAAAC 1080  
Db 1395 ATCATGGAAGAGCAAACTCTGTACCGAAAGTATCTGAGGATGCTCTCGAAATGTG 1454  
Qy 1081 AACGCATTTGATAGTGGCTGTCAAGAGATCTTCTCGAGAGGTGAGTATGCTATA 1140  
Db 1455 GATGCTTTGACAAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGAGTATGCACTC 1514  
Qy 1141 AAGTATCCGTGGCATAGAGTATGCCAAGATTGGAGGCAAGAGTTACATAGACAAATTT 1200  
Db 1515 AAATATCCCTGGCATAGAGTATGCCAAGTTGGAGGCTAGAAGCTATATTTGAAGAACTAT 1574  
Qy 1201 GGATCAAAATGATGCTGGCTGGGGAAGACTGTGTATAGATGCTATATGTGAGCAACGAA 1260  
Db 1575 GGGCCAGATGATGCTGGCTGGGAAACCTGTATATATGATGATGATGATGATGATGATGAT 1634  
Qy 1261 AATATTTGGAGCTGGGCAATGGAATTCATATGTTGCGAGGCTTTACACCAAAAGGAG 1320  
Db 1635 AAGTATTTAGAACTAGCGAACTGGAATTCATTAAGGTGCACTATACACCAAAACAGAG 1694  
Qy 1321 ACTCAACACATTTGCTGAGTGGTGGAGAAATCGGATTCATGATCTTACATTTACCCGC 1380  
Db 1695 CTTCAAGATCTTCAAGGTGGTGAATTCATCGGTTTCAAGGATCTGAAATTTCACTCGT 1754  
Qy 1381 CAGCGGCTGTGGAATGATTTCTCAGTGGGCTTAGTATGTTGAGCCAGATTCGCT 1440  
Db 1755 GAGGCTGTGCGGAATATATTTCTCACCGGATCTTTATCTTTGAGCCGAGTTTCT 1814  
Qy 1441 GCTTGTAGATGCTGATGCAAGACTTCTGCTCGAGATTTCTGAGAGATCTTTTAC 1500  
Db 1815 AAGTGCAGAGAGGTTTATACAAAATCTTCAATTTCACTGTTATTTTAGATGATCTTTAT 1874  
Qy 1501 GACACCAACGATCTGAGTATCTTAATATGTTCTTCAAGGCTTCGGAAGATGGAT 1560  
Db 1875 GAGCCCATGATCTTTAGACGATCTTAAGTTGTTTACAGAAATCAGTCAAAAGATGGAT 1934  
Qy 1561 ATCTCTGCTGATAGCTTTCGGGATAATCAGTTGAAAGTTTCTTCTTAGGCTCTGAC 1620  
Db 1935 CTATCACTAGTGA---CCAAATGCCACAAATGAATATATGTTTGGGTTTCTAC 1991  
Qy 1621 AACACAGTGAATGGAATTTGGAAGATGGAATCAAGGAACAAGCCGCTGATGTCGGC 1680  
Db 1992 AATACTTTAATGATATAGCAAAAGAGAGCTGAGAGGCAAGGCGGATGTCGATGCTAGGC 2051  
Qy 1681 TATCTTGAAGATGAGGAGGCTTCTGCTGATGATATACCAAGAGCGGAATGTCG 1740  
Db 2052 TACATTCAAAATGTTTGAAGATCAACTTTGAAGCTTACAGAAAGAGCAGATGCTCT 2111  
Qy 1741 GCAGCAAGTATGTCGCAATTCACGAATATGTAAGAAATGCAAGTGTCCATAGCA 1800  
Db 2112 GAAGCTAAATGTTGGCAATCTTCAATGAATACATAGAGATGCGATGTTGTCATAGCA 2171  
Qy 1801 CTTGCGACGTCGTAATCAATCTTTTCACTGGAGAAATTTACTTCTCGATTAAT 1860  
Db 2172 TTGGGAACAGTCGTTCTCATTTAGTCTTTTCACTGGGAGGTTCTTACAGATGAAGTA 2231  
Qy 1861 TTACAGCAAGTAGACCTTCGTCGATCAATTTCTGATCTTGTCTTTTGACTGGAGACTA 1920  
Db 2232 CTCTCCAAAATGATCGCAATCTAGATTTCTTCAACTCATGCGCTTTAAACAGGCGGTTG 2291  
Qy 1921 ATCAATGACCAAGACTTACAGCGCGAGAGAAACGTCGTGATGTTTTCAGCGTA 1980  
Db 2292 GTGAATGACCAAAACTTTATCAGGACAGAGAGGTCAAGGTGAGGTGGCTTCTTGCCATA 2351

Qy 1981 CAGTGTACATGATGAGGGAATTCGGAGTGCACAGAGAGAGAGCTCTAAGTATGTTTAT 2040  
Db 2352 CAATGTTTATGAGGACCATCTCTAAATCTCTGAAGAGAGAGCTCTACAACTGTCTAT 2411  
Qy 2041 GGTATCATCCACAAACGCACTGAAGGAATTCGAATTTGGGAGTTGGCAACCCAGCGAGCAAT 2100  
Db 2412 AGTGTATGGAATATGCCCTCGAAGAGTTGAATAGGAGTTTGGAA-----TACAAA 2455  
Qy 2101 GCCCAATTTGTGTGAGAGAGACTGCTGTTTCAACACTCAAGAGTGTGAGCTGTTTAT 2160  
Db 2466 ATACCGGATATTTACAAAAGACTGGTTTTTGAAGCTCAAGAAATTAATGCAACTCTTTAT 2525  
Qy 2161 ATGTACAGAGATGCTTTGTTGATCTC---TGACAAAGAGATGAAAGACCATGTGAGCGGA 2217  
Db 2526 ATGCAAGGGATGTTTGAACACTATCATATGTAATTTAAAGAGCATGTCAAAAAT 2585  
Qy 2218 ACTCTTTTCCATCTCTGGCGTAG 2241  
Db 2586 TGCTCTTTCCAAACAGTTGCCTAG 2609

## RESULT 9

US-09-900-797-55  
; Sequence 55, Application US/09900797  
; Publication No. US20030087406A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20030087406A1, Joseph P.  
; APPLICANT: Starke, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 55  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3)...(2606)  
; OTHER INFORMATION: abietadiene synthase  
US-09-900-797-55

Query Match 48.1%; Score 1078; DB 10; Length 2861;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

Qy 1 ATGTTTTCAGTCCATGGCGGACGGTGAACCGAATCCATCTGATATGATACAGCTTGGTG 60  
Db 378 ATGTTTAGATGATGGCTATGGGAAACGAAATCCCTCTGATATGACATGCTTGGTA 437  
Qy 61 GCAAGAAATCCGTCATTTGACGCTCTGTGCAACCCCAATTTCCCAAAACGCTTCAATGG 120  
Db 438 GCAAGGATTCAGCAGTTGATGCTCTGACAAACCTTCTCTGAGAGCGTTGAATGG 497  
Qy 121 ATTCTGAACAACTCAATGCCAGATGCTCTGGGTGAGGAGTGCATTTTCTGGCGTAT 180  
Db 498 ATTCTTCAAAATCAGTTGAAGATGGGTCTTGGGGTGAAGGATTTCTACTTCTTGGCATAT 557  
Qy 181 GACAGAGTTTTAAACACTCTCGGCTGCTCTCACTCTCAAAATATGGAATTAAGGGGAC 240  
Db 558 GACAGATATCGCTACACTTGCATGATGATTTATACCTTACCTCTGCGGTACTGGGGAG 617

QY	241	ATTCAAGTGCAGAAAGGGTTCAGTTTGTGAGAAACACATCGGAAGAAATGAAGACGAA	300	Db	1695	CTTCAAGATCTTCGAAGGTGTGGAAATCATCCGGTTTACCGATCTGAATTTCACTCGT	1754
Db	618	ACACAAGTACAGAAAGGTATTGAATTTCTTCAGGACACAAAGCTGGAAAGATGGAAGATGA	677	QY	1381	CAGCGGCTGTGGAATGATTTCTCAGTGGCGGTTAGTATGTTTGGAGCCAGAAATTCGCT	1440
QY	301	GCTGACAAATCACAGGCCAAGTGGATTCGAGGTGCTGTTTCTTCGCAATGTTAGATGAAGCA	360	Db	1755	GAGCGTGTGACGGAATATATTTCTCACCAGCATCTTTATCTTTGAGCCCGAGTTTCT	1814
Db	678	GCTGATAGTCATAGGCCAAGTGGATTTGAAATAGTATTTCTTCGCAATGCTAAAGGAAGCT	737	QY	1441	GCTTGTAGAATTGCTATGCCAAGACTTCTTGCTCCGACTTATTTAGACGATCTTTTAC	1500
QY	361	AAAAGCTTGGGATTTGATCTTCTTATCACCTCCCTTTTCATCTCCCAATTCACCAAAAG	420	Db	1815	AAGTGCAGAGAGGTTTATACAAAACCTTCCAAATTTCACTGTTATTTAGATGATCTTTAT	1874
Db	738	AAAATCTTAGGCTTGGATCTGCTTACGATTTGCCATTTCTGAAACAAATCATCGAAAG	797	QY	1501	GACACCCACGAGTCTCTGGATGATCTTAAATTTGTTCTCTGAAGCGTCCGAGATGGAT	1560
QY	421	CGCCAGAAAAGCTTCAAAAGATTTCCCTCAATGTTTCTTCAATACCATCAGACGGGTTG	480	Db	1875	GAGCCCATGGATCTTTAGACGATCTTAAAGTTGTTTACAGAAATCAGTCAAAAGATGGAT	1934
Db	798	CGGAGGCTAAGCTTAAAGGATTTCCCACTGATGTTCTCTATGCTCTTCCAAACAGTTA	857	QY	1561	ATCTCTGTCTGGATAGCGTTTCGGGATAATCAGTTGAAAGTTTGTCTTCTTAGGCTGTAC	1620
QY	481	CTCTACTCTGAGGGTTTGCAGATGTTGGTGGACTGGCAAGAGATCAAAATCTTCAA	540	Db	1935	CTATCACTAGTGGG---CCAAATGCCACAACAATGAAATATGTTTGTGGGTTTCTAC	1991
Db	858	TTGTATTTCTTGAAGGTTTACAGAAATAGTAGACTGGCAGAAATTAATGAACCTTCAA	917	QY	1621	AACACAGTGAATGGATTTGGAAAGATGGACTCAAGGAAACAAGCCGTGATGTCTGGGC	1680
QY	541	TCAGAGACGATCATTTTAAAGCTCCCTGCACTACTGCTTGTGCTTTCATGCAACT	600	Db	1992	AATACITTTTATGATATAGCAAAAGAGGACGTGAGAGGCAAGGGCGGATGTCTAGGC	2051
Db	918	TCCAGAGATGATCATTTCTCAGCTCTCCGCACTACAGCGGCTGATTCATGCGGTACA	977	QY	1681	TATCTTCGAAAAGTATGGAGGCTTGTCTCGCATCTGTATACAAAGAAAGTGTCCATAGCA	1740
QY	601	CAAAACAAACGATGCCCTTCTCACTTCTCGTCTCAGCAAAATTCGCGCACTACGTT	660	Db	2052	TACATTTCAAAATGTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGAGCAGATGTCT	2111
Db	978	GGGAACAAAGATGCTTGGATTTCTTGAACCTTGTCTTGAAGAAATTCGGAACCATGTG	1037	QY	1741	GCAGCAAAAGTATGTCCGACATTCACGAAATATGTGGAAATGTCGAAAGTGTCCATAGCA	1800
QY	661	CTTTGCCATTCACCACTTGATCTATTTGAAGCGCTCTGGGCTGTCGATACAGTTGAACGC	720	Db	2112	GAAGCTAAATATGTGCCATCTTCAATGAATACATAGAAATGCGAGTGTGTCAATAGCA	2171
Db	1038	CTTTGTCACTATCCGCTTGATCTATTTGAAGCTTGTGGGCGTTGATACAGTTGAGCGG	1097	QY	1801	CTTSCGACAGTCTGTAATAACCTCACTTTTCTACTGGAGAAATTTACTTCTGATTACAT	1860
QY	721	TTGGGAATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGATACGTTTATAGG	780	Db	2172	TTGGGAACAGTCTGTTCTCTATTTAGTGTCTTTTCTACTGGGAGGTTCTTACAGATGAAGTA	2231
Db	1098	CTAGTATCGATCGTCAITTTCAAGAGGAGATCAAGGAAGCATTTGGATTTATTTTACAGC	1157	QY	1861	TTACAGCAAGTATGACCTTCGGTCCAAATTTCTGCAATCTTGTCTTTTGACTGGAGCACTA	1920
QY	781	TACTGGGACGCGAAGAGGCGTGGATGGGCAAGATGCAATCTTATTCCTGATGTCGAT	840	Db	2232	CTCTCCAAAATTTGATTCGCGAATCTAGATTTCTTCAACTCATGGGCTTAAACAGGGCGTTTG	2291
Db	1158	CAATTGGGA---CGAAGAGGCAATTTGATGGGCGAGAGAGAAATCTCTGTTCTGATATTGAT	1214	QY	1921	ATCAATGACACCAAGACTTTACAGGCGGAGAGAAACCGTGTGATTTGGTTTCCAGCGTA	1980
QY	841	GACATGCGATCGGCTTAGAATCTCGAGACTTCAATGGATACAAATGATCTTTCAGATGTT	900	Db	2292	GTGAATGACACCAAACTTTATCAGGCGAGAGAGAGGTCAAGGTGAGGTGGCTTCTGCCATA	2351
Db	1215	GATACAGCCATGGGCTTCCGATCTTTGATTTACATGATGATCAATTTGATCTCTCAGATGTT	1274	QY	1981	CAGTGTCTACATGAGGAAATCCGAGTGCACAGAGAAAGAAAGCTCTTAAGTCAATGTTTAT	2040
QY	901	CTGGGAATTTCAAGACGAGAAAGGAGACTTCTTTGCTTTGCGGTCAAACGCAAAAT	960	Db	2352	CAATGTTTATATGAAGGACCATCTTAAATCTCTGTAAGAAAGAGCTCTACAACTGTCTAT	2411
Db	1275	TTAAAAACATTTAGAGATGAGAAATGGGAGTTCTTTTGTCTTCTTGGGTCAAAACAGAGA	1334	QY	2041	GGTATCATCGACAAACGCACTGAAGGAATTTGAATTTGGGAGTTGGCCCAACCCAGCGAGCAAT	2100
QY	961	GGTGTGACCGTAATCTTAACTTTATAGATGTTCAAAAGTATGTTTTCGGGAGAAAG	1020	Db	2412	AGTGTCAATGGAATATGCTCTGTAAGAGAGTTGAATAGGAGGTTTGTGAA-----TAACAA	2465
Db	1335	GGAGTTACAGACATGTTAAACGTCAATCGTTGTTTCAATGTTTCAATTTCCGGGAGAAAG	1394	QY	2101	GCCCCATTTGTGTGAGAGACTGCTGTTTCAACACTGCAAGAGTGTATGAGCTGTTTAT	2160
QY	1021	ATAATGGAAGAGCTAAGACCTTCACTAACAATCATCTCCAAATGCTCTTCCCAAAAC	1080	Db	2466	ATACCGGATATTTACAAAGAGCTGGTTTTTGAACCTGCAAGAAATATGCAAGTAATATGCAACTCTTTAT	2525
Db	1395	ATCATGGAAGAGCAAAACTCTGTACCGAAAGGATCTGAGGAATGCTCTGGAATATGTG	1454	QY	2161	ATGTACAGAGATGCTTGTGGTATCTC---TGCAAGAGATGAAAGACCATGTCCAGCGGA	2217
QY	1081	AACGCAATTTGATTAAGTGGGCTGTCAAGAGGATCTTCTTGGAGAGGTGGAGTATGCTATA	1140	Db	2526	ATGCAAGGAGTGTGTTTGACACTATCACATGATATGGAATTAAGAGCATGTCAAAAT	2585
Db	1455	GATGCGCTTTGACAAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGATATGCACTC	1514	QY	2218	ACTCTTTTCGATCTGTGGCGTAG	2241
QY	1141	AAGTATCCGTCGATAGAAAGTATGCCAAGATTTGGAGCAAGAGTTTACATAGAGCAATTT	1200	Db	2586	TGCCTCTTCCAAACAGTTGCGCTAG	2609
Db	1515	AAATATCCCTGGCATAGAGTATGCCAAGTTGGAGGCTAGAAAGCTATATTTGAAACTAT	1574				
QY	1201	GGATCAAAATGATGTCTGGCTGGGAGACATGTTGATTAAGATGCTATATGTGAGCAACGAA	1260				
Db	1575	GGGCGCATGATGTGTGGCTTGGAAACACTGTATATATGATGCCATACATTTTCAATGAA	1634				
QY	1261	AAATATTTGAGCTGGCCAAATTTGGAATTCATATGTTGTCAGGCTTTTACACCAAGAGG	1320				
Db	1635	AAGTATTTAGAACTAGCGAAACTGGACTTCAATAAGGTGAGTCTATACACCAACAGAG	1694				
QY	1321	ACTCAACACATTTGTACGCTGTGGAGAGAAATCGGGATTCATGATCTTACATTTCCACCGC	1380				

RESULT 10

US-10-041-365

; Sequence 365, Application US/10041018

; Publication No. US2004007232A1

; GENERAL INFORMATION:

; APPLICANT: Matusda, Seichi P.T.

; APPLICANT: Hart, Elizabeth A.

; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism

; FILE REFERENCE: P020800SI/10025547

RESULT 10  
 US-10-041-018-365  
 ; Sequence 365, Application US/10041018  
 ; Publication No. US2004007232A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Matsuda, Seiichi P.T.  
 ; APPLICANT: Hart, Elizabeth A.  
 ; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
 ; FILE REFERENCE: P02080U51/10025547

; CURRENT APPLICATION NUMBER: US/10/041.018  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259880  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 365  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
US-10-041-018-365

Query Match 48.1%; Score 1078; DB 12; Length 2861;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

QY 1 ATGTTTCAGTCCATGGCGCAGCGGTGAAACGAATCCATCTGCATATGATACAGCTTGGGTG 50  
DB 378 ATGTTTAGATGTATGGCTATGGCGAATCCCTCTGCATATGACACTGCTTGGTA 437  
QY 61 GCAAGAAATCCGTCATATGACGGCTCTGGTGCACCCCAATTTCCGCCAAGCTTCAATGG 120  
DB 438 GCAAGGAATTCAGCAGTTGATGGCTCTGCAACAACCTCCTCTCTGAGACGGTTGAATGG 497  
QY 121 ATTCTGAACAATCAACTGCCAGATGGCTCGTGGGTGAGGAGTGCAATTTTCTGGCGTAT 180  
DB 498 ATTCTTCAAAATCAGTTGAAAGATGGGTCTTGGGGTGAAGGATTCATCTCTTGGCATAT 557  
QY 181 GACAGAGTTTAAACACTCTCGCTCGCTCCTCACTCTCAAAATATGGAATAAGGGCGAC 240  
DB 558 GACAGATACTGGCTACACTTGATGATTAATTAACCTTACCCTCTGGCGTACTGGGGAG 517  
QY 241 ATTCAAGTGCAGAAAGGGTTGATTTGTGAGAAAAACACATGGAAGAAATGAAGACGAA 300  
DB 618 ACACAAGTACAGAAAGGATTTGAATTTCTCAGGACACAAGCTGGAAGATGGAAGTAA 577  
QY 301 GCTGACATACAGGCCAAGTGAATTCGAGGTGCTGTTTCTGCAATTTAGATGAAGCA 360  
DB 678 GCTGATAGTCATAGGCCAAGTGAATTTGAATATAGTATTTCTGCAATGCTTAAAGGAAGCT 737  
QY 361 AAAAGCTTGGGATTTGATCTTCTTATCACTCCCTCTTCACTCCCAATCCCAACAAAG 420  
DB 738 AAAATCTTAGGCTTGGATCTGCTTACGATTTGGCATCTCTGAAACAAATCATCGAAG 797  
QY 421 CGCCAGAAAAAGCTTCAAAAGATTCGCCCTCAATTTGTTCTTCAATCAACCAATCAGCGGGTTG 480  
DB 798 CGGAGGCTAAGCTTAAAGGATTCACCATGATGTTCTATGCCCTTCCAAACAGTTA 357  
QY 481 CTCTACTCTCGAGGGTTTGAAGATGTTGGTGGACTGGCAAGAGATCAAAATCTTCAA 340  
DB 858 TTGTATTTCTTGAAGGTTTACAGAAATAGTAGACTGGCAGAAAAATGAATGAAACTTCAA 917  
QY 541 TCAAGAGACGGATCATTTTAAAGCTCCCTCGCATCTACTCTGTGTCTTTCATGCACT 500  
DB 918 TCAAGATGATCATTTCTCAGCTCTCGGCACTTACAGCGCTGTATTTATGCGGTACA 377  
QY 601 CAAAAAAGAGTGCCTTCACTTTCTCAACTCGTGTCTCAGCAAAATTTGCGCACTACGTT 560  
DB 978 GGGAAACAAAAAGTCTTGGATTTCTTGAACATTTGTCTTGAAGAAATTCGGAACCATGTG 1037  
QY 661 CTTTGCATTAACCATTTGATCTATTTGAACGCTCTGGCTGTGCATACAGTTGAACGC 720  
DB 1038 CTTTGTCACTATCCGCTTGAATCTATTTGAACGTTTGTGGCGGTTGATACAGTTGAGCGG 1097  
QY 721 TTGGGAATCGATCGTATTTTCAAGAAAGAAATCAAGAAATCTCTGGATACGTTTATAGG 780  
DB 1098 CTAGGTATCGATCGTCAATTTCAAGAGAGATCAAGAGCATTTGATATGTTTACAGC 1157  
QY 781 TACTGGGACGCCGAAAGAGCGGTGGATGGGCAAGATGCAATCTTATCTCTGATGTCGAT 340  
DB 1158 CATTTGGGA---CGAAGAGGCAATGGATGGCGGAGAGAAATCTGTTCTCTGATATGAT 1214  
QY 841 GACACTGCCATGGGTCTTAGAATCTTGAGACTTTCATGGATACAAATGATCTTCAGATGTT 900

DB 1215 GATACAGCCATGGCCCTTGGAACTTTGAGATTAATGATACAAATGATATCTCAGATGTT 1274  
QY 901 CTGGAGAAATTTAGAGACGAGAAAGGAGACTCTCTTTTGGCTTTGCGGCTCAAAACCAAAAT 960  
DB 1275 TTAANAACATTTAGAGATGAGATGGGAGTCTCTTTTGGCTTTCTGGCTCAAAACACAGAGA 1334  
QY 961 GGTGTGACCGATTAATCTTAACTTTATAGATGTTTCAAGATGTTTTCGGGGAGAAAAG 1020  
DB 1335 GGAGTTACAGACATGTTAAACGTCATCGTTGTTTCAATGTTTCAATTTCCGGGAGAAACG 1394  
QY 1021 ATAAATGGAAGAGCTAAGACCTTCACTACAAAATCATCTCCAAAATGCTCTTGGCCAAAAC 1080  
DB 1395 ATCATGGAAGAGCAAAACCTCTGTACGAAAGTATCTGAGGAATGCTCTGGAAAATGTTG 1454  
QY 1081 AACGCATTTGATTAAGTGGGCTGTCAAGAAGGATCTTCTCTGGAGAGTGGAGTATGCTATA 1140  
DB 1455 GATGCTTTTGACAAATGGGCTTTTAAAGAAATATTCGGGGAGAGGTAGAGTATGCATC 1514  
QY 1141 AAGTATCCGTGGCATAGAGTATGCCAAGATTTGGAGCAAGAAAGTTACATAGAGCAATTT 1200  
DB 1515 AAATATCCCTGGCATTAAGAGTATGCCAAGGTTGGAGGCTAGAAGCTATATTTGAAAAC 1574  
QY 1201 GGATCAAAATGATCTCTGGCTGGGGAAGACATGCTGATAAGATGCTATATGAGCAACGAA 1260  
DB 1575 GGCCAGATGATGTGGCTTGGAAAACCTGTATATATGATGCCATACATTTTGGAAATGAA 1634  
QY 1261 AAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGTTGTCAGGCTTTTACACCAAAAGGAG 1320  
DB 1635 AAGTATTTAGAACTAGGGAACCTGGACTTCAATTAAGGTGCAGTCTATACACCAAAAGAG 1694  
QY 1321 ACTCAACATTTGTCAGCTGGTGGAGAGAAATCGGATTCGAATCAATGATCTTACATTCACCCG 1380  
DB 1695 CTTCAAGATCTTCAAGGTTGGTGAATCATCCGGTTTACCGGATCTGAAATTTCACTCGT 1754  
QY 1381 CAGCGGCTGTGAAATGATTTCTCAGTGGCGGTAGTATGTTTGAAGCCAGAAATTCGCT 1440  
DB 1755 GAGCGGTGTCAGGAAATATATTTCTCACCGGATCTCTTATCTTGAAGCCGAGTTTCT 1814  
QY 1441 GCTTGTAGAAATTTGCCATGCAAGACTTCTTGCTCGCAGTATTTCTAGACGATCTTTAC 1500  
DB 1815 AAGTCAGAGAGGTTTATACAAAACCTTCAATTTCACTGTTATTTAGATGATCTTTAT 1874  
QY 1501 GACACCCAGGATCTCTGGATGATCTTAAATTTCTCTGAAGCGGTTCGGAAGATGGAT 1560  
DB 1875 GACGCCATGATCTTTAGACGATCTTAAAGTTCTTACAGAAATCAGTCAAAAGATGGAT 1934  
QY 1561 ATCTGTGCTGGATAGCTTGGGATTAATCAGTTGAAAGTGTGCTTCTTAGGCTGTAC 1620  
DB 1935 CTATCACTAGTGA---CCAAATGCCACAAATGAAATATGTTTGTGGGTTTCTAC 1991  
QY 1621 AACACAGTCAATGGAATTTGGAAAAGATGGAATCAAGGCAAGGCGGTGATGTGCTGGGC 1680  
DB 1992 AATACCTTTAATGATATAGCAAAAGAGGACGTCGAGGCAAGGCGGATGTGCTAGGC 2051  
QY 1681 TATCTTGGAAAAGTATGGAGGCTTGTGTCGATCGTATACCAAGAGCCGAATGCTG 1740  
DB 2052 TACATTTCAAAATGTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGAGCAGAAATGCT 2111  
QY 1741 GCAGCAAGTATGTGCGGACATTTCAAGCAATATGTCGAAAATGCCAAAGTTCATAGCA 1800  
DB 2112 GAAGCTTAAATATGTGCGCATCTCTTCAATGAATACATAGAGAAATCGAGTGTGCTCAATAGCA 2171  
QY 1801 CTTGCGACAGTCTGATCAAACTCAATCTTTTCACTGGAGAAATTAATCTCTGATTAAT 1860  
DB 2172 TTGGGAAGAGTCTCTTCAATAGTCTCTTCACTGGGAGGTTCTTACAGATGAAGTA 2231  
QY 1861 TTACAGCAAGTAGACCTTGGTCCAAATTTCTGCAATTTCTGCTTTGATGTCGACGACTA 1920  
DB 2232 CTCTCCAAATTTGATCGCAATCTAGATTTCTTCAACTCATGGGCTTAAACAGGCGGTTG 2291  
QY 1921 ATCAATGACCAAGACTTACAGGCGGAGAGAAACGGTGGTGAATTTGGTTCCAGCGTA 1980



Thu Aug 26 08:27:54 2004

```

Db      2292 GTGAATGACACCAAACTTATCAGGCAGAGAGAGGTCAAGGTGAGTGGCTTCTGCCATA 2351
Qy      1981 CAGTGTACATGAGGAAATCCGGAGTCACAGAGAGAGAGTCTTAAGTCAATGTTTAT 2040
Db      2352 CAATGTTATATGAGGACCACTCTAAATCTCTGAAGAAGAGCTCTACAATGTCAT 2411
Qy      2041 GGTATCATCGACAAACGCACTGAGAGAAATGAATTTGGAGTTGGCCACCAGCGAGCAAT 2100
Db      2412 AGTGTATGGAATATGCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAAA 2465
Qy      2101 GCCCATTTGTTGAGAGAGCTGTGTTCAACACTGCAAGAGTGTGAGCTGAGCTGTTTAT 2160
Db      2466 ATACCGGATATTACAAAAGACTGGTTTTTGAACACTGCAAGATAATGCAACTCTTTAT 2525
Qy      2161 ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCAATGTGACCGA 2217
Db      2526 ATGCAAGGGAGTGGTTTGACACTATACATGATATGGAATTAAGAGCATGTCAAAAT 2585
Qy      2218 ACTCTTTTCGATCTGTGGCGTAG 2241
Db      2586 TGCTCTTCCAAACAGTTGCTAG 2609

RESULT 11
US-09-893-820-55
; Sequence 55, Application US/09893820
; Publication No. US20040053386A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20040053386A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/893,820
; PRIOR FILING DATE: 2001-05-28
; PRIOR APPLICATION NUMBER: US/09/398,395A
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 55
; LENGTH: 2861
; TYPE: DNA
; ORGANISM: Abies grandis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(2606)
; OTHER INFORMATION: abietadiene synthase
US-09-893-820-55

Query Match      48.1%; Score 1078; DB 13; Length 2861;
Best Local Similarity 68.8%; Pred. No. 0;
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;

Qy      1 ATGTTTCAGTCCATGGCGGACGGTGAACGAATCCATCTGCATATGATACAGCTTGGGTG 60
Db      378 ATGTTTAGATGATAGGCTATGCGAAACGAATCCCTCTGCAATGACACTGCTGGGTA 437
Qy      61 GCAAGAAATCCCGTCAATGAGCGCTGTGTGCAACCCCAATTTCCCAAGCTTCAATGG 120
Db      438 GCAAGAAATCCAGCAGTTGATGGCTCTGACAAACCTTCTCTGAGACGGTTGATGG 497
Qy      121 ATTCTGACAAATCAACTGCGAGATGGCTCGTGGGGTGAGGAGTGCATTTTCTGGCGTAT 180
Db      498 ATTCTTCAAAATCAATGTTGAAGATGGGTCTTGGGGTGAAGGATCTACTTCTTGGCATAT 557
Qy      181 GACAGAGTTTTAAACACTCTGCGCTGCTCTCACTCTCAAAATATGGAATAGGGCGAC 240

```

```

Db      558 GACAGAACTAGCTGGCTACACTTCATGTTATTATTACCTTACCTCTGGCGTACTGGGAG 617
Qy      241 ATTCAAGTCAGAAAGGGTTCAGTTTGTGAGAAACACATGGAAGAAATGAAGACGAA 300
Db      618 ACACAAATGACAGAAAGGTATTGAATTTCTTCAGGACAAAGCTGGAAGATGAAGATGA 677
Qy      301 GCTGCAATTCACAGGCAAGTGGATTTCGAGGTGCTGTGTTTCTGCAATGTTAGATGAAGCA 360
Db      678 GCTGATAGTCATAGGCAAGTGGATTGAAATAGTATTTCCTGCAATGCTAAAGGAGCT 737
Qy      361 AAAAGCTTCGGATTTGATCTTTCATCACTCCCTTTCATCTCCCAATCCACCAAG 420
Db      738 AAAATCTTAGCTTGGATCTGCTTACGATTTGCCATTCCTGAAACAAATCATCGAAG 797
Qy      421 CGGCAGAAAGCTTCAAAAGATTTCCTCAATGTTCTTCATACCAATCCACAGCGGTG 480
Db      798 CGGAGGCTTAAGCTTAAAGATTTCCTCAATGTTCTTCATGCTTCCCAACAACTGA 857
Qy      481 CTCTACTCTCTGGAGGGTTTTCAGAGATGTGGTGGACTGGCAAGAGATCAAAATCTTCAA 540
Db      858 TTGTTATTTCTTGAAGGTTTACAAAGAAATAGTAGACTGGCAGAAATAATGAATCTTCAA 917
Qy      541 TCAAGAGACGGATCATTTTAAAGCTCCCTGCTACTACTCTGTTGCTTTCATCACTCA 600
Db      918 TCCAAGGATGGATCATTTCTCAGCTCTCCGCTCTACAGGGCTGTATTCATCGGTACA 977
Qy      601 CAAACAAACGATGCTCCACTTTCTCAACTTCTGCTCTGAGCAAAATTTGGGCACTACGTT 660
Db      978 GGGAAACAAAGTCTGGATTCTTGAACCTTCTTGAAGAAATTCGGAACCAATGTG 1037
Qy      661 CTTTGGCCATTACCCACTGTATTTAAAGCCCTCTGGCTGTGATACAGTTGAAAGCG 720
Db      1038 CTTTGTCACTATCCGCTTGTATTTGAAGCTTGTGGGGGGTGTGATACAGTTGAGCGG 1097
Qy      721 TTGGAAATCGATCGCTATTTCAAGAAAGAAATCAAGAAATCTCTGGAATGATCTTATAGG 780
Db      1098 CTAGGTATCGATCGTCAATTTCAAGAGGAGATCAAGGAAGCATTTGATTTATACAGC 1157
Qy      781 TACTGGGACGCGGAAAGGGGTGGGATGGGCAAGATGCAATCTTATTCCTGATCCGAT 840
Db      1158 CATTTGGGA---CGAAAGAGGCAATGGATGGGCGAGAGAAATCTCTGTTCTGATATGAT 1214
Qy      841 GACTGCTCATGGCTCTTAGATCTGAGACTTCAATGATACAAATGATATCTTTCAGATGTT 900
Db      1215 GATACAGCCATGGGCTTTCGAAATCTTGAGATTTACATGATACATGATATCTTCCAGATGT 1274
Qy      901 CTGGGAATTTTCAGAGACGAGAAAGGAGACTTTTGTGCTTGGCGGTCAACGCAAAAT 960
Db      1275 TTAATAACATTTAGAGATGAGAAATGGGAGTTCTTTGCTTCTGGGTCAACACACAGAG 1334
Qy      961 GGTGTGACCGATAATCTTAACTTTATAGATGTTTCAAGATGTTTCCGGGAGAGAAAG 1020
Db      1335 GGAGTTACAGACATGTTAAACGTCATCGTGTGTCAGATGTTTCAATTCGCGGAGAGAACG 1394
Qy      1021 ATAATGGAGAGAGCTAAGACCTTTCATCAAAATCATCTCAAAATGCTCTTGCACAAAC 1080
Db      1395 ATCATGGAAAGAGCAAAACTCTGTACCGAAAGTATCTGAGGAATGCTCTGGAATAATGTG 1454
Qy      1081 AACGCAATTTGATAAGTGGGCTGTCAAGAGGATCTTCTGAGAGAGTGGAGTATGCTATA 1140
Db      1455 GATGCTTTGACAAATGGGCTTTTAAAGAAATATTTCCGGGAGAGGTAGATGATGCACTC 1514
Qy      1141 AGATATCCGTGGCATAGAGATGCAAGATTCGAGGCAAGAGTTTACATAGAGCAATTT 1200
Db      1515 AAATATCCCTGGCATAGAGATGCAAGGTTGGAGGCTAGAGCTATATTTGAAACTAT 1574
Qy      1201 GGATCAAAATGATCTGTGGCTGGGGAAGACTGTGTATAAGATGCTATATGTAGCAACGAA 1260
Db      1575 GGCCAGATGATCTGTGGCTTGGAAACCTGTATATATGATGCTCATATTTTCAATTTGAA 1634
Qy      1261 AAATATTTGAGCTGGCCAAATTTGGACTTCAATATGCTGAGGCTTACACCAAAAGGAG 1320

```

Db 1635 AAGTAATTAGAACTAGCGAACTGGACTTCAATAAGGTGAGTCTATACACCAACACAGAG 1694  
QY 1321 ACTCAACACATTTGCTAGCTGGTGGAGAGATCGGGATTCAATGATCTTAATTCACCCGC 1380  
Db 1695 CTTCAAGATCTTGAAGTGGTGGAAATCATCCGGTTTCACGGATTTGAATTTCACTCGT 1754  
QY 1381 CAGCGGCTGTGGAATGTAATTTCTCAGTGGCGTTAGTATGTTTCAGCCAGAAATTCGT 1440  
Db 1755 GAGCGTGTGACGGAATATATTTCTCACCGGCATCTTTATCTTTGAGCCGAGTTTCT 1814  
QY 1441 GCTTGTAGAAATGCTATGCAAGACTTCTTGCTCGCAGTATTTCTAGACGATCTTTAC 1500  
Db 1815 AAGTCAGAGAGGTTTATACAAAATCTTCAATTTCACTGTTATTTAGATGATCTTAT 1874  
QY 1501 GACACCCACGATCTCTGGATGATCTTAAATGTTCTCAAGCGGTCGGAAGATGGAT 1560  
Db 1875 GAGCCCATGGATCTTTAGACGATCTTAAGTTGTTACAGAAATCAGTCAAAAGATGGAT 1934  
QY 1561 ATCTCTGCTGTAGTAGCGTTCCGGATTAATCAGTTGAAAGTTTGTCTTCTAGGCTCTAC 1620  
Db 1935 CTATCACTAGTGA---CCAAATGCCCAACAAATGAAATATGTTTGTGGGTTTCTAC 1991  
QY 1621 AACACAGTGAATGGATTTGAAAGAGTGGACTCAAGGAAACAGGCGGTGATGTCTGGG 1680  
Db 1992 AATACTTTTATGATATAGCAAAAGAGAGCTGAGAGGCAAGGCGGATGTCTAGG 2051  
QY 1681 TATCTTCGAAAGATATGGGAGGCTGCTCGCATCGTATACCAAGAACGCGAATGGTGG 1740  
Db 2052 TACATTCAAAATGTTTGGAAAGTCCAACCTTGAAGCTTTACACGAAAGAGCAGAATGGTCT 2111  
QY 1741 GCAGCAAGTATGTCGCAATTCACCGATATGTGGAATATGCGAAATGCCAAGTCCATAGCA 1800  
Db 2112 GAAGCTAAATATGTGCCATCTTCAATGAATATACATAGAGAAATCGAGTGTGTCAATAGCA 2171  
QY 1801 CTTGCGACAGTCTGCTATAACTCAATCTTTTCACTGGGAAATTAATCTCTGATTAAT 1860  
Db 2172 TTGGGAACAGTCTGTTCTATAGTCTCTTTCACTGGGAGGTTCTTACAGATGAAT 2231  
QY 1861 TTAACAGAGTAGACCTTCGGTCCAAATTTCTGCATCTTGTCTTTGACTGGACGACTA 1920  
Db 2232 CTCTCCAAATTTGATCGGAATCTAGATTTCTTCACTATGGCTTTAAACAGGCGGTTG 2291  
QY 1921 ATCAATGACACCAAGACTTACAGGCGGAGAGAAACGCTGTGTAATGGTTTCCAGGTA 1980  
Db 2292 GTGAATGACACCAAACTTATCAGGACAGAGAGGTTCAAGGTGAGGTGGCTTCTGCCATA 2351  
QY 1981 CAGTGTCTACATGAGGAAATTCGGAGTGACAGAGGAAGCTCTAAGTCAATGTTAT 2040  
Db 2352 CAATGTTATGAGGACCACTCTTAATCTCTGAAGAGAGCTCTAACAATGCTAT 2411  
QY 2041 GGTATCATCGACCAACGCACTGAAGGAATGAATTTGGGAGTTGGCCAAACCCAGCGAGCAAT 2100  
Db 2412 AGTGTCAATGAAATGCCCTCGAAGAGTTGAATAGGAGTTTGTGAA-----TAACAA 2465  
QY 2101 GCCCATTTGTGTGAGAGACTGCTGTTCAACACTGCAAGAGTATGAGCTGTTTAT 2160  
Db 2466 ATACCGGATATTTACAAAGACTGGTTTTTGAACCTGCAAGAAATAATGCAACTCTTTAT 2525  
QY 2161 ATGTACAGAGATGGCTTTGGTATCTC---TGACAAAGAGATGAAGACCAATGTCCAGCGA 2217  
Db 2526 ATGCAAGGGATGGTTTGACACTATCATGATATGGAATTAAGAGCAATGTCAAAAT 2585  
QY 2218 ACTTTTTCGATCTGTGGCGTAG 2241  
Db 2586 TGCCTCTTCCACACAGTTGCTAG 2609

## RESULT 12

US-10-041-007-38  
; Sequence 3, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.

; APPLICANT: Schepmann, Hala G  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081051  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; PRIOR FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 3  
; LENGTH: 2861  
; TYPE: DNA  
; ORGANISM: Abies grandis  
US-10-041-007-38

Query Match 48.1%; Score 1078; DB 14; Length 2861;  
Best Local Similarity 68.8%; Pred. No. 0;  
Matches 1544; Conservative 0; Mismatches 685; Indels 15; Gaps 4;  
QY 1 ATGTTTCAGTCCATATGGCGGACGGTGAAACGAATCCATCTCGCATATGATACAGCTTGGGTG 60  
Db 378 ATGTTTAGATGATATGGGCTATGGCGAAACGAATCCCTCTGCATATGACACTGCTTGGGTA 437  
QY 61 GCAAGAAATCCCGTCAATTTGACGCTCTGTGTCACCCCAATTTCCCAAAACGCTTCAATGG 120  
Db 438 GCAAGGATTCACAGCTGATGGCTCTGACACCCCTCCTTTCTGAGCGTTGAATGG 497  
QY 121 ATTCTGAAACAATCAATGCCAGATGGCTCGTGGGGTGAGGAGTGCAATTTTTCTGGCGTAT 180  
Db 498 ATTCTTCAAAATCAGTTGAAAGATGGTCTTGGGGTGAAGGATTTCTACTTCTTGGCATAT 557  
QY 181 GACAGAGTTTAAACAATCTCGCTCGCTCTCACTCTCAAAATATGGAATAGAGGCGAC 240  
Db 558 GACAGAAATCTGGCTACACTTGCATGATTAATTAACCTTACCCTCTGCGCTACTGCGGAG 617  
QY 241 ATTCTAGTGCAGAAAGGGTTGAGTTTGTGAGAAACACATGGAAGAAATGAAGAGCA 300  
Db 618 ACAAGTACAGAAAGGATTTGAATTTCTCAGGACACAGCTGGAAGATGGAAGATGAA 677  
QY 301 GCTGCAAAATCACAAGGCAAGTGGATTCGAGGTGCTGTTTCTTGCATATGTTAGATGAAGCA 360  
Db 678 GCTGATAGTATAGGCAAGTGGATTTGAAATAGTATTTCTTGCATATGCTAAAGGAAGCT 737  
QY 361 AAAAGCTGGGATTTGATCTTCTTATCACTCCCTTTCATCTCCCAATTCACCAAAAG 420  
Db 738 AAAATCTTAGGCTTGGATCTGCTTACGATTTGCCATTTCCCTGAAACAAATCATCGAAAG 797  
QY 421 CGCCAGAAAGGCTTCAAAAGATTTCCCTCAATGTTCTTCAATCAATCAGACGGCGTTG 480  
Db 798 CGGAGGCTAAAGCTTAAAGGATTTCCCATGATGTTCTTATGCCCCCTTCCAAACAGTTA 857  
QY 481 CTCTACTCTCTGAGGGTTTGAAGATGTGGTGGACTGGCAAGAGATCACAATCTTCAA 540  
Db 858 TTGATTTCTTGAAGGTTTACAGAAATAGTAGACTGGCAGAAATATGAAACTTCAA 917  
QY 541 TCAAGAGACGATCAATTTTAACTCCCTGCAATCTAATCTGCTGTGTCTTCAATGCACT 600  
Db 918 TCCAAGATGATCAATTTCTCAGCTCTCCGCACTTACAGCGGCTGATTAATCATGCGTACA 977  
QY 601 CAAAACAAAGGCTCCACTTTCTCACTCTGCTGCTCAGCAATTTGGCGACTACGTT 660  
Db 978 GGAACAAAAGTGTCTTGAATTTCTTGAATTTGCTTTGAAGAAATTCGGAACCATGTG 1037  
QY 661 CTTTGGCAATTAACCCACTTGAATCTTATTTGAACGCTCTTGGGCTGTGCAATAGTGAACGC 720  
Db 1038 CTTTGTCACTATCCGCTTGAATCTTATTTGAACGTTTGTGGCGGTTGATACAGTTGACGG 1097  
QY 721 TTGGGAATCGATCGCTATTTTCAAGAAAGAAATCAAGAAATCTCTGGAATAGTTTATAGG 780  
Db 1098 CTAGGTATCGATCGTCAATTTCAAGAGGAGATCAAGGAAGCAATGGAATATGTTTACAGC 1157  
QY 781 TACTGGGACGCCAAAGAGCGGTGGGATGGGCAAGATGCAATCTTATTTCTCTGATGTCAT 840

1158 CATTTGGGA---CGAAGAGGCATTGGATGGCGGAGAGAGAAATCCTGTTCTGTATATTGAT 1214  
841 GACACTGGCCATGGTCTTAGAATCCTGAGACTTTCATGGATACAAATGATATCTTCAGATGTT 900  
1215 GATACAGCCATGGCCCTTCGAAATCTTGAGATTACATGGATACAAATGATCTTCAGATGTT 1274  
901 CTGGAGAAATTTACAGACGAGAAAGAGACTTCTTTTGGCTTTCCGGTCAACGCAAAAT 960  
1275 TTAABAACTTTAGAGATGAGATGGGAGTCTTTTGGCTTTCCGGTCAACGCAAGAG 1334  
961 GGTGTGACCAATATCTTAACCTTTATAGATGTTCAAGATGTTTTCGGGAGAAAG 1020  
1335 GGAGTTACACACATGTTAAACGTCATCTGTTTCAATGTTTCAATTTCCGGGAGAAAG 1394  
1021 ATAATGGAAGAGCTTAAGACCTTCACATACAAATCATCTCCAAATGCTCTTGGCAAAAC 1080  
1395 ATCATGGAAGAGCAAACTCTGACCGAAGGATCTGAGGAATGCTCTGGAATGTTG 1454  
1081 AAGCATTTGATAGTGGCTGTCAAGAGATCTTCTCGAGAGGTGGAGTATGCTATA 1140  
1455 GATGCTTTGACAAATGGGCTTTTAAAGAAATATTCGGGAGAGGTAGAGTATGCACTC 1514  
1141 AAGTATCCGTGGCATAGAGATGATCCCAAGATTCGAGGCAAGAGTTCATATAGAGCAATTT 1200  
1515 AATATCCCTGGCATAGAGATGATCCCAAGGTTGAGGCTAGAGCTATATTTGAAATCTAT 1574  
1201 GGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATATAGATGCTATATGAGCAACGA 1260  
1575 GGGCCAGATGATGTCTGGCTTTGGAATACTGTATATATGATGATGATGATGATGATGATG 1634  
1261 AATATTTGGAGCTGGCCAAATTTGAGTCTCAATATGTTGAGGCTTACACCAAAAGGAG 1320  
1635 AAGTATTTAGAACTAGCGAACTGAGACTTCAATAGTTCATAGTTCATATACCAAAAGAG 1694  
1321 ACTCAACATTTGTCAGCTGGTGGAGAGAAATCGGATTCATATGATCTTACATTCACCCG 1380  
1695 CTTCAAGATCTTCGAAGGTGGTGGAAATCATCCGGTTTCAAGGATCTGAAATTTCACTCG 1754  
1381 CAGCGGCTGTGGAAATGATTTCTCAGTGGCGGTAGTATGTTTGGCCAGAAATTCGCT 1440  
1755 GAGCGTGTGACGGAATATATTTCTCACCGGATCCTTTATCTTTTGGCCGAGTTTCT 1814  
1441 GCTTGTAGAAATGCTATGCCAAGACTTCTTCCCTCGCAGTTATTTCTAGAGATCTTTAC 1500  
1815 AAGTCAGAGAGTTTATACAAAACTTCCATTTTCACTGTTATTTTATAGATGATCTTTAT 1874  
1501 GACACCCAGGATCTCTGATGATCTTAAATGTTCTCTGAAGCGTCCGAGATGGGAT 1560  
1875 GACGCCCATGGATCTTTAGACGATCTTAAATGTTTTCACAGAACTCAGTCAAAAGATGGAT 1934  
1561 ATCTGTGCTGATAGCGTTCCGGATAATCAGTTGAAAGTTTGTCTTCTAGGCTGTATC 1620  
1935 CTATCACTAGTGA---CCAAATGCCACAAATGAAATATGTTTGTGGGTTTCTAC 1991  
1621 AACACAGTGAATGGATTTGGAAAGATGAGCTCAAGAAACAGGCCGCTGATGCTGGG 1680  
1992 AATACTTTTAAATGATATAGCAAAAGAGCGTGGAGGCAAGGCGGCGATGCTAGGC 2051  
1681 TATCTTGAAGATGATGGAGGCTTGTCTGCAATGATATACCAAGAAAGCGAATGGTCTG 1740  
2052 TACATTTCAAAATGTTTGGAAAGTCCAACTTGAAGCTTACACGAAAGAGCAAGATGGTCT 2111  
1741 GCAGCAAGATGATGCGCAATTTCAACGAATATGTGGAAATGCCAAAGTGTCCATAGCA 1800  
2112 GAAGCTAATATGTGCACTCTTCAATGAATACATAGAAATGCGAGTGTGTCAATAGCA 2171  
1801 CTTGGGACAGTGTACTTAATCAATCTTTTCACTGGAAATTTACTTCTGATTTACAT 1860  
2172 TTGGGAACAGTGTCTCTCATTTAGTGTCTTTTCACTGGGAGGTTCTTACAGATGAGTA 2231  
1861 TTACAGCAAGTGTAGACCTTCCGTTCCAAATTTCTGATCTTGTGCTTTTGTGCTGAGCA 1920  
2232 CTCTCCAAATTTGATCGGATCTAGATTTCTTCAACTCATGGCTTAAACAGGCGGTTG 2291

1921 ATCAATGACCAAGACTTTACAGGCGGAGAGAAACCGTGGTGAATTTGGTTTCCAGCGTA 1980  
2292 GTGAATGACACCAAACTTTATCAGGCGAGAGAGGTCGAAGTGGTGGCTTCTGCCATA 2351  
1981 CAGTGTACATGAGGAAATCCGAGTGCACAGAGGAAGAGCTCTAAGTCAATGTTTAT 2040  
2352 CAATGTTATATGAGGACCATCTTAAATCTCTGAAGAGAGCTCTTACCAATGTTCTAT 2411  
2041 GGTATCATCGACCAACCACTGAGGAAATGAATTTGGAGTTGGCAACCCAGCGAGCAAT 2100  
2412 AGTGTATGAGAAATGCCCTCGAAGAGTTGAAATAGGAGTTTGAA-----TAACAA 2465  
2101 GCCCATTTGTGTGAGAGAGACTGCTGTTCAACATCTGCAAGAGTGTAGTGTGTTTAT 2160  
2466 ATACCCGATTTTACAAAGACTGTTTGAACACTGCAAGAAATATGCAACTCTTTTAT 2525  
2161 ATGTACAGAGATGGCTTTGTTATCTC---TGACAAAGAGATGAAGACCAGTGTGCGGA 2217  
2526 ATGCAAGGGATGGTTTGCACACTATCATATGATATGGAATTAAGAGCATGTCAAAAT 2585  
2218 ACTCTTTTCGATCTCTGGGCTAG 2241  
2586 TGCCTCTTCCACAGTTGCTAG 2609

RESULT 13  
US-09-887-586A-45  
; Sequence 45, Application US/09887586A  
; Patent No. US20020094556A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: NO. US20020094556A1, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/887,586A  
; PRIOR FILING DATE: 2001-06-22  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(2347)  
; OTHER INFORMATION: E-alpha-bisabolene synthase  
US-09-887-586A-45

Query Match 32.3%; Score 723.4; DB 9; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 6.8e-222;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

31 AATCCATCTGCATATGATACAGCTTGGGTGGCAAGAAATTCGTCATTTGAGTGGCTCGT 90  
134 ACTCCATCTGCTTATGACACAGCATGGGTAGCGAGGTTGCCCGCATTTGATGGCTCTGCT 193  
91 GCACCCCAATTTCCCAAGAGCTTCAATGATTTCTGAACAACTCAACTGCCAGATGGCTCG 150  
194 CGCCCGCAATTTCCCAAGAGCTTGAATTTGAAACCAAGTTAAAGATGGTTCA 253  
151 TGGGGTGAGGAGTGCATTTTCTTGGGCTATGACAGAGTTTAAACACTCTCGCTGCTC 210  
254 TGGGGGAATTCAGTCCCACTTTCTGCTCGACCGTCTTCTTGCCACTCTTTCTTGTGTT 313

Qy	211	CTCACTCTCAAAATATATGGAATTAAGGGCGACATTCACGTGCAGAAAGGGGTGAGTTTGTG	270
Db	314	CTTGTGCTCTCTAAATGCGAAGCTTTGGGGATCTGCAAGTAGAGCAGGGAAATTGAATTCATA	373
Qy	271	AGAAAAACACATGGAGAAATGAGGACGAGCTGACAAATCACAG--GCCAAGTGGATTC	327
Db	374	AAGAGCAATCTGGAACTAGTAAGAGATGAACCGATCAAGATAGCTTGGTAAACAGACTTT	433
Qy	328	GAGTCTGTCTTCTCTGCAATCTTGTAGATGAAGCAAAAAAGCTTTGGGATTTGGATCTTCTCTTAT	387
Db	434	GAGATCATATTTCTCTCTCTGTTAAGAGAAGCTCAATCTCTGCGCTCGGACNCTTCCCTAC	493
Qy	388	CACCTCCCTTTCATCTCCCAAAATCCACAAAGGCCAGAAAAAGCTTCAAAAGATTTCCC	447
Db	494	GACCTGCTTATATACATCTGTTTGCAGATAAACCGCAGGAAAGATTTAGCAAAACTTTCA	553
Qy	448	CTCAATGTTCTTCATAACCATCAGACGGCGTGTCTCTACTCTCTGAGGGGTTTGCCAAGAT	507
Db	554	AGGAGGAAAAATTTATGCGGTTCCGTCGCCAATTTGTGTATTTCTTTAGAGGGAAATACAAGAT	513
Qy	508	GTGTGTGACTGGCAAGAGATCACAAATCTTCAATCAAGAGACGGATCATTTTTTAAGCTCC	567
Db	614	ATAGTTGAATGGGAAACGAATAATGGAAGTTCAAAGTCAGATGGGTCCTTTCTTAAGCTCA	573
Qy	568	CTGTGATCTACTGCTGTGCTTTCATGCAACATCAAAACAAACGATGCCCTCCACTTTCTC	527
Db	674	CCTGCTTCTACTGCTCGTCTTTTCATGCAACACAGGAGACGCGAAATGCCCTTGAATCTCTG	733
Qy	628	AACCTTGTGCTCAGCAAAATTTGGGACTACGTTCTTGGCCATTTACCACTTTGATCTTATTT	587
Db	734	AACAGTGTGATGATCAAGTTTGGGAAATTTTGTCTCCCTGCTGATCTCTGTGGATCTGCTG	793
Qy	688	GAACGCTCTTGGGCTGTGATACAGTTTGAACGCTTTGGGAATCGATGCTATTTCAAGAAA	747
Db	794	GAACGCTCTTGTGATCGTAGATATATTTGTAGCCTTTGGAATCTATAGACACTTTGAAAG	353
Qy	748	GAATCAAGAATCTCTGGATTAGTTTATAGGTACTGGGACGCCGAAAGAGCGCTGGGA	307
Db	854	GAATCAAGGAAGCTCTTGATATGTTTACAGGCAATTGGAA---CGAAACAGGAAATTGGG	310
Qy	808	TGGGCAAGATGCAATCTCTATTTCTGATGTGATGAGACATCGCCATGGGTCTTAGAATCCTG	367
Db	911	TGGGGCAGACTAAATCCCATAGCAGATCTTTGAGACCACTGCTTTTGGGATTTTCGATTCCT	370
Qy	868	AGACTTCATCGATACAAATGTATCTTCAGATGTTCTTGAGAGATTTTCAGAGACGAGAAGGA	327
Db	971	CGGCTGATAGGTACAAATGATCTCCAGCCATTTTGCACACTTTCAAGATGCCATGGG	1030
Qy	928	GACTTCTTTTGTCTTTCGGTCCAAACGCAAAATTTGGTGTGACCGGATATCTTAACCTTTAT	387
Db	1031	AAATTCATTTGCTCGACGGTCAATTCAAACAAGATGTAGCAAGCATGCTGAATCTTTAT	1090
Qy	988	AGATGTTCAACAATGATTTTTTCCGGGAGAAAAAGATAATGGAAGAGCTAAGACCTTCACCT	1047
Db	1091	AGAGCTTCCAGCTCGCATTTTCCCGGAGAAAACAATCTTGATGAAGCTAAAGAGCTTCGCT	1150
Qy	1048	ACAAATCATCTCBAATGCTCTTGCCAAAAACAACGCAATTTGATAGTGGGCTGTCAAG	1107
Db	1151	ACTAAATATTTGAGAGAAGCTCTTTGAAAAAGTGAGACTTCCAGTGCATGGAACAACAA	1210
Qy	1108	AAGGATCTTCTCGAGAGGTGGAGTAGTGTATAAAGTATCCGTGGCATAGAAGTATGCCA	1167
Db	1211	CAAAAACCTGAGCCAAAGATACAAATACGCCCTGNAAGCTTCTTGGCATGCCAGTGTCCG	1270
Qy	1168	AGATTGGAGCAAGAAGTTACATAGAGCAATTTGGATCAAATGATGTCTGGCTGGGGAAG	1227
Db	1271	AGAGTGGAAAGCAAGAGATACGTGTCAAGTGTATCGCCAGATTTATGCACGCATAGCAAAA	1330
Qy	1228	ACTGTGTATAGATGCTATATGTGAGCAACGAAAAATATTTGGAGCTGGGCCAATTTGGAC	1287
Db	1331	TGGGTTTACAGCTPACCTACGTGGAACAAATGAAAAAGTTTTTAGAGCTGGGAAATTTAGAT	1390
Qy	1288	TTCAATATGCTGAGGCGCTTACACCAAAAAGGAGACTCAAACATTTGTACGCTGTGGAGA	1347

1391	Db	 TTCAAATTATCCAGTCCATCCACCAAGAAGAAATGAAGAAATGTTACCAAGCTGGTTTAGA
1348	Qy	 GAATCGGGATCAATGATCTTACATTCACCCGCCAGCGCCCTGTGGAATGATATTTCTCA
1451	Db	 GATTCGGGGTTGCCATATTCACCTTCGCTCGGGAGAGCCGCTGGAAATCTCTACTTTCTTA
1408	Qy	 GTGCGGTTAGTATGTTTGAGCCAGAAATTCGCTGCTGTAGAAATGCGCTATGCAAGACT
1511	Db	 GTAGCGCGGGACCTATGNAACCCAGTATGCCAAATGCAGGTTCTCTTACAAAGTG
1468	Qy	 TCCTTGCTCGCAGTTATCTTAGACGATCTTTACGACACCCACCGGATCTCTGGATGATCTT
1571	Db	 GCATGCTTGACAGACTGTTCTGGACGATATGTTGACACATTAATGGAACCCCTAGATGAATTG
1528	Qy	 AAATTTGTTCTCTGAAGCGGTCCGAAGATGGGATATCTCTGTGCTGGATAGGCTTCGGAT
1631	Db	 AAGCTATTCACTGAGGCTGTGGAAGATGGGACCTCTCTCTTACAGAAAACCTTCAGAC
1588	Qy	 AATCAGTTGAAAGTTGCTTCCTPAGGCTGTCAACACACAGTGAATGGAATTCGAAAAGAT
1691	Db	 TAT---ATGAAACTATGTTACCAAAATCTATATATGACATAGTTCCACAGGTGGCTTCGGAG
1648	Qy	 GGACTCAAGNAACAAGCCGTGATGTCTGGGCTATCTTCGAAAAGTATGGGAGGCTTG
1748	Db	 GCAGAGNAAGNACAAGGGCGTGATTTGGTCAGCTTTTTCAGAAAAGGATGGGAGGATTAAT
1708	Qy	 CTCGCATCGTATACCAAGAAGCCGAATGGTCGGCAGCAAAGTATGTGCCGACATTCAC
1808	Db	 CTTCTGGGTTATTTATGAAGAAGCTGAATGGTTAGCTGTGAGTATGTGCTACCTCTGGAC
1768	Qy	 GAATATGTGGAAATGCCAAAGTGTCATAGCNCCTTGGCAGCAGTCGTACTTAACTCAATC
1868	Db	 GAGTACATAAAGAAATCGAAATCACATCTATCGGCCAACGTATATCTCTGTTGAGTGAGTG
1828	Qy	 TTTTTTCAC---TGGAGAAATTTACTTCTCGATTAATCATTTTACAGCAAGTACAGCTTCGGTCC
1928	Db	 TTGATTAATGATGGGCACATCTCTTCGAAAGGCAATTAGAAGAAAGTATGATTAATCAGGA
1885	Qy	 AAATTTCTGC-----ATCTTGTCCTTTTGACTGGACGACTAATCAATGACACCAAGACT
1988	Db	 AGACGTGTTCTCACAGAGCTGAATAGCCTCATTTCCGCTTCGCGGATCACAGAGACA
1939	Qy	 TACACGCCGAGAGAAACCGTGGTGAAATTTGGTTTCAGCGTACAGTGCTACATGAGGAA
2048	Db	 TATAAAGCTCAGAAGGCTCGTGAGAAATTTGGGCTCCAGCATTTGAAATGTTTACATGAAGAC
1999	Qy	 AATTCGGGATGACACAGAGNAGACTCTAAGTCATGTTTATGTTATCATTCGACAACGCA
2108	Db	 CATCTCTGAAATGATACAGAGAGAGGCTCTCGATTCACATCTATAGCATTCCTGGAGCGGG
2059	Qy	 CTGAAGGAATGGAATTTGGGAGTTGGCCAACCCAGCGAGCAATGCCCCATTTGTGTGTAGA
2168	Db	 GTGAAGGAATGACAGAGAGTTTCTGAAGCCG---ACGAGTCCCATTCGCTGCGCAG
2119	Qy	 AGACTGCTGTTCAACACTGCAAGAGTGATGCAGCTGTTTTTATATGATCAGAGATGGCTTT
2225	Db	 AAGATGCTTTTCGAGGACACAAGAGTGACGATGGTGATATTCAAGGATCGAGATGGATTTC
2179	Qy	 GGTATCTCTGACAAAGAGATGAAGACCATTGTCCAGCCGAACTCTTTTTCGATCTCTGTGC
2285	Db	 GGTGTTTTCAAAATTAGAAGTCAAAGATCATATCAAAGAGTGCTCATTCGAAACGCTGCC

RESULT 14  
US-09-903-012-45  
; Sequence 45, Application US/09903012  
; Patent No. US20020094557A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20020094557A1, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.

Thu Aug 26 08:27:54 2004

;; TITLE OF INVENTION: SYNTHASES  
;; FILE REFERENCE: 07678-025001  
;; CURRENT APPLICATION NUMBER: US/09/903,012  
;; CURRENT FILING DATE: 2001-07-11  
;; PRIOR APPLICATION NUMBER: 09/398,395  
;; PRIOR FILING DATE: 1999-09-17  
;; PRIOR APPLICATION NUMBER: 60/100,993  
;; PRIOR FILING DATE: 1998-09-18  
;; PRIOR APPLICATION NUMBER: 60/130,628  
;; PRIOR FILING DATE: 1999-04-22  
;; PRIOR APPLICATION NUMBER: 60/150,262  
;; PRIOR FILING DATE: 1999-08-23  
;; NUMBER OF SEQ ID NOS: 58  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 45  
;; LENGTH: 2424  
;; TYPE: DNA  
;; ORGANISM: Abies grandis  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (2)...(2347)  
;; OTHER INFORMATION: E-alpha-bisabolene synthase  
US-09-903-012-45

Query Match 32.3%; Score 723.4; DB 9; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 6.8e-222; Indels 21; Gaps 6;  
Matches 1327; Conservative 0; Mismatches 871;

QY	31	AATCCATCTGATATGATACAGCTTTGGGTGGCAAGAAATTCGGTCAATTTGACGGCTCTGGT	50
DB	134	ACTCCATCTGTTATGACACAGCATGGGTAGCGGGTGGCCGCCAATTTGATGGCTCTGCT	193
QY	91	GCACCCCAATTTCCCAAAACGCTTCATGGATTTCTGAACATCACTGCCAGCGGCTCG	150
DB	194	CGCCGCAATTTCCCAAAACGTTGACTGGATTTTGAACAAACCACTTAAAGATGGTTCA	253
QY	151	TGGGTGAGGAGTGATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCTGCGCTC	210
DB	254	TGGGAATTCAGTCCACTTTCTGCTGTCGACCGCTCTTTGCGCACCTTTCTTGTGT	313
QY	211	CTCACTCTCAAAATATGAATGAAGGCGACATTTCAAGTGCAGAAAGGGTTGAGTTGTG	270
DB	314	CTTGTGCTCTTAAATGAAACGTTGGGATCTGCAAGTATGACGAGGAAATTTGAAATTCATA	373
QY	271	AGAAACACATGGAAGAAATGAAGCAAGCTGACAACTACAG--GCCAAGTGGATTC	327
DB	374	AAGAGCAATCTGGAACCTAGTAAGGATGAAACCGATCAAGATAGCTTGTACAGACTTT	433
QY	328	GAGGTGCTGTTTCTTCGCAATGTATAGTGAAGCAAAAGCTTGGGATTTGGATCTTCTTAT	387
DB	434	GAGATCAATTTCTCTCTCTGTTAAGAGAACTCAATCTCTGCGCTCGGACTTCCCTAC	493
QY	388	CACCTCCCTTTCATCTCCCAATCCCAAGCGCCAGAAAGCTTCAAAAGATTCC	447
DB	494	GACCTGCTTATATACATCTGTTGCAAGCTAAACCGCAGAAAGATGACAAACTTCA	553
QY	448	CTCAATGTTCTTCATAACCATCAGACGGGTGTCTACTCTCTGAGGGTTTGAAGAT	507
DB	554	AGGAGGAAATTTATCGGTTTCGTCGCCATTTGTTGTTTCTTTAGGGAATACAAAT	613
QY	508	GTGGTGGACTGGCAAGAGATCAAAATCTTCAATCAAGAGACGGATCATTTTAAGCTCC	567
DB	614	ATAGTTGAATGGGAACGAATATGGAAGTTTCAAAGTCAGGATGGGTCTTTCTTAAGCTCA	673
QY	568	CTTGCACTACTGCTTGTCTCTCATGCACTCAAAACAAACGATGCGCTTCCACTTTCTC	627
DB	674	CTGCTTCTACTGCTGCTTTTTCATGCAACAGGAGACGGGAATGCTTGAATCTTG	733
QY	628	AACCTCGTCTCAGCAAAATTTGGCGACTACGTTCTTGTGCAATTTACCACTTTGATTT	687
DB	734	AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCGCTGTATCTGTGGATCTGCTG	793
QY	688	GAAAGCTCTGGGCTGTCGATACAGTGTGAACGCTTGGGAATCGATCGCTATTTCAAGAA	747

DB	794	GAACGCTGTGATCGTAGATAATATTGTACGCTTGGAAATCTATAGACTTTGAAAG	853
QY	748	GAATCAAGAAATCTCTGGATTTAGTACTTGGGACCGCAAGAGCGTGGGA	807
DB	854	GAATCAAGGAAGCTCTTGAATATTTTACAGCATTTGAA--CGAAGAGGAATTTGGG	910
QY	808	TGGGCAAGATGCAATCTATTTCTGATGTGATGACACTGCCATGGCTTTAGATCTTG	867
DB	911	TGGGGCAGACTAATCCCATAGCAGATCTTGAGACCACTGCTTTGGGATTTGATTCCT	970
QY	868	AGACTTCATGGATCAATGTATCTTCAATGTTCTGGAGAAATTTACAGACGCAAGGA	927
DB	971	CGGCTGCATAGTACAAATGTATCTCCAGCCATTTTTCACAACTTTCAAGATGCAATGGG	1030
QY	928	GACTTCTTTTGTGCTTGGCGGTCAAAACGCAAAATTTGGTGTGACCGGATATTTAACTTTAT	987
DB	1031	AAATTCATTTGCTCGACCGGTCAATTTCAACAAAGATGTAGCAAGCATGCTGAATCTTTAT	1090
QY	988	AGATGTTTCAAGTATGTTTTCCGGGAGAAAGTAATGGAAGAACTTAAGACTTCACT	1047
DB	1091	AGAGCTTCCAGCTCGCATTTTCCCGGAGAAACATTTCTTGATGAAGCTTAAAGCTTCCCT	1150
QY	1048	ACAAATCATCTCAAAATGCTTTTCCCAAAACAAACGCAATTTGATAGTGGGCTGTCAAG	1107
DB	1151	ACTAAATATTTGAGAGAGCTTTTGAGAAAGTGAAGCTTCCAGTGCATGGAACAAACA	1210
QY	1108	AAGGATCTTCTGAGAGAGTGGAGTATGCTATAAAGTATCCGTGGCATAGAAATGCA	1167
DB	1211	CAAACTGAGCAAGAGATCAAAATACGCGTGAAGACTTCTTGGCATGCCAGTGTCCG	1270
QY	1168	AGATTGAGGAGAAAGATTACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGAAG	1227
DB	1271	AGAGTGGAAAGCAAGAGATCTGTCAAGTGTATGCCCGAATTTAGCAGCATACAAA	1330
QY	1228	ACTGTATAGAGTCTATATGTGAGCAACGAAATAATTTGGAGCTGCCCAATTTGAC	1287
DB	1331	TGCGTTTACAGCTACCTACGTGAACAATGAAAGTTTTTAGAGCTGGGAAATTTAGAT	1390
QY	1288	TTCAATATGGTGGAGCCTTACACAAAGAGAGCTCAACATTTGTCAGCTGGTGGAGA	1347
DB	1391	TTCAACATTTCCAGTCCATCCACCAAGAGAAATGGAAGATGTTACCAAGCTGGTTTGA	1450
QY	1348	GAATCGGGAATCAATGATTTTACATTTCCCGCCAGCGCTGTGGAAATGATTTTCTCA	1407
DB	1451	GATTCGGGGTGGCCTACTATTTACCTTCTCGCTCGGAGAGGCCGCTGGAATTTACTTCTTA	1510
QY	1408	GTGGCGGTTAGTATGTTTGAAGCCAGAAATTCGCTGCTGTGAGAAATTCCTATGCCAGACT	1467
DB	1511	GTAGCGGGGAGCCTATGAACCCAGTATGCAAAATGCAGGTTCTCTTTTCAAAAGTG	1570
QY	1468	TCTTGGCTCGAGTTATTTTAGACGATCTTTACGACACCCAGGATCTCTGATGATCTT	1527
DB	1571	GCATGCTTGACAGCTGTTCTGACGATATGATGACACTTNTTGAACCTTAGATGATTTG	1630
QY	1528	AAATTTGTTCTGAAAGCGTCCGAAAGATGGGATATCTCTGCTGGATAGCGTTCGGGAT	1587
DB	1631	AGACTATTCATGAGGCTGTGAGAGATGAGACCTCTCTCTTACAGAAACCTTCCAGAC	1690
QY	1588	AATCAGTTGAAAGTTGCTTCTTAGGGCTGTACACACAGTGAATGATTTGGAAGAT	1647
DB	1691	TAT--ATGAAACTATGTTACCAAAATCTATTATGACATAGTTCCAGAGTGGCTTGGGAG	1747
QY	1648	GGACTCAAGGAACAAGCGCTGATGCTGGGCTATCTTTCGAAAGATGATGGGAGGCTTG	1707
DB	1748	GCAGAGAGGAACAGGGGCTGAAATTTGCTGAGCTTTTTCAGAAAGGATGGAGGATTTAT	1807
QY	1708	CTCGCATCGTATACAAAGAGCCGAATGCTCGGAGAGCAAGATGATGTGCCGACATTTCAAC	1767
DB	1808	CTTCTGGTTTATGTAAGAGAGCTGAATGGTTAGTGTGCTGAGTATGTGCTTACCTTGAC	1867
QY	1768	GAATGTGGAAATGCCAAAGTGTCCATAGACTTTCGAGCAGTGTGCTACTTAACTCAATC	1827

Db 1868 GAGTACATGAAGATGAATCACAATCTATCGGCCAACGTATATCTTCTGTTGAGTGGAGTG 1927  
Qy 1828 TTTTTCAC----TGAGAAATTAATCTCTGATATACATTTTACAGCAAGTAGACCTTCGGTCC 1884  
Db 1928 TTGATATGATGGCACTCTTTCCGAAGAGGCATTAAGAGAAAGTAGATTAATCCAGGA 1987  
Qy 1885 AAATTTCTGC-----ATCTTGTTGCTTTTCTGACTGGAGCACTAATCAATGACACCAAGACT 1938  
Db 1988 AGACGTGTTCTCAGAGCTGAATAGCCTCATTTTCCCGCTGGCGGATGACAGAGACA 2047  
Qy 1939 TACCAGCCGAGAGAAACCGTGGTGAATTTGTTTCCAGCGCTACAGTCTACATGAGGAA 1998  
Db 2048 TATAAGCTGAGAGGCTCTGGAGAAATTTGGCTCCAGCATTTGATGTTACATGAAGAC 2107  
Qy 1999 AATCCGAGTGACAGAGGAAGAGCTTAAGTCAATGTTTATGATATCATGCAACGCA 2058  
Db 2108 CATCTGAATGTACAGAGGAAGAGGCTCTCGATCACAATCTATAGCAATCTGGAGCCGCG 2167  
Qy 2059 CTGAAGGAATGAATTTGGGAGTTGGCCAAACCGAGGCAATGCCCATTTGTGTGAGA 2118  
Db 2168 GTGAAGGAATCTGAAGAGAGTTTCTGAAGCCCG---ACGACGTCCCATTTGGCCTGCAAG 2224  
Qy 2119 AGACTGCTTTCAACACTCGAAGAGTATGAGCTGTTTATATGTATGATCAGAGATGGCTTT 2178  
Db 2225 AAGATGCTTTTCAGGAGACAAGTGCAGATGCTGATATTCAGGATGGAGATGGATTC 2284  
Qy 2179 GGTATCTCTGACAAAGAGATGAAGACCATGTGACCGAACTCTTTTCGATCTCTGGC 2237  
Db 2285 GGTGTTTCCAAATTTAGAAGTCAAGATCATATCAAGAGTGTCTCATTTGAACCGCTGCC 2343

## RESULT 15

US-09-900-797-45  
; Sequence 45, Application US/09900797  
; Publication No. US20030087406A1  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. US20030087406A1, Joseph P.  
; APPLICANT: Starke, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/900,797  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: US/09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PastSeq for Windows Version 3.0  
; SEQ ID NO 45  
; LENGTH: 2424  
; TYPE: DNA  
; ORGANISM: Abies grandis  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(2347)  
; OTHER INFORMATION: B-alpha-bisabolene synthase  
US-09-900-797-45

Query Match 32.3%; Score 723.4; DB 10; Length 2424;  
Best Local Similarity 59.8%; Pred. No. 6.8e-222;  
Matches 1327; Conservative 0; Mismatches 871; Indels 21; Gaps 6;

Qy 31 AATCATCTGATATGATACAGCTTGGGTGGGCAAGAAATTCGTCAATTTGACGGCTCTGGT 90  
Db 134 ACTCCATCTGTTATGACACAGCATGGGTAGCGAGGGTGGCCGCCAATTTGATGGCTCTGCT 193  
Qy 91 GCACCCCAATTTCCCAAAACGCTTCAATGGATTTCTGAACATCAACTGACCATGGCTCG 150  
Db 194 CGCCCGCAATTTTCCCAAAACAGTTGACTGGATTTTGAAGAAACCAAGTTTAAAGATGTTCA 253

Qy 151 TGGGTTGAGAGATGCAATTTTCTGGCGTATGACAGAGTTTAAACACTCTCGCCTGCCTC 210  
Db 254 TGGGAAATTCAGTCCCACTTTCTGCTGTCGCGCGTCTTCTTGCCTCTTCTTGTGTT 313  
Qy 211 CTCACCTCTCAAAATATGGAATTAAGGCGACATTTCAAGTGCAGAAAGGGTTGAGTTGTG 270  
Db 314 CTTGTGCTCTTAAATGGAACGTTTGGGATCTCGAAGTAGAGCGGAAATGAAATTCATA 373  
Qy 271 AGAAAACACATGGAAGAAATGAAGGACGAAAGCTGACAAATCACAG---GCCAAGTGGATTC 327  
Db 374 AAGAGCAATCTGGAACCTAGTAAAGGATGAACCGATCAAGATAGCTTTGGTAAACAGACTTT 433  
Qy 328 GAGGTGCTGTTTCTGCAATGTTAGATGAAGCAAAAGCTTTGGATTTGGATTTCTTCTTAT 387  
Db 434 GAGATCATATTTCTCTCTGTTAAGAGAAAGCTCAATCTCTGCGCCTCGGACTTTCCCTAC 493  
Qy 388 CACTCTCTCTCTCTCTCCAAATCCCAAAAGCGCCAGAAAAGCTTTCAAAAGATTTCCC 447  
Db 494 GACTGCTTATATATACATCTGTTGACAGCTAAACGGCAGGAAGATTAAGCAAACTTTCA 553  
Qy 448 CTCATGTTTCTTCAATACCATCAGACGGCGTTGCTCTACTCTCTGAGGGTTTGCAGAT 507  
Db 554 AGGAGGAAATTTATGCGGTTCCGTCGCCATTTGTTGTTATCTTTAGAGGGAATACAAGAT 613  
Qy 508 GTGGTGAATGCGCAAGAGATCAAAATCTTCAATCAAGACGAGATCATTTTAAAGCTCC 567  
Db 614 ATAGTTGAATGGGAACGGAATTAATGGAAGTTTCAAGTCAGGATGGGTCTTTCTTAAAGCTCA 673  
Qy 568 CTGCGATCTACTGCTGTTGCTTCTATGSCACACTCAAAACAAACGATGSCCTCCACTTTCTC 627  
Db 674 CTTGCTTCTACTGCTGCTGTTTCTATGACACAGGAGACCGGAAATGCTTTGAATTTCTTG 733  
Qy 628 AACTTCTGCTCAGCAAAATTTGGCGACTAGCTTCTTTCCTTTCCTTACCACTTTGATCTATTT 687  
Db 734 AACAGTGTGATGATCAAGTTTGGAAATTTTGTTCCTGCTGCTGATCTCTGTTGGATCTGCTG 793  
Qy 688 GAAAGCTCTGGGCTGTGATACAGTTGAAGCTTGGGAATCGATTCGCTATTTTCAAGAAA 747  
Db 794 GAAAGCTCTGTGATCGTAGATAATAATGTAAGCTTGGAAATCTATAGACACTTTTGAAGAA 853  
Qy 748 GAAATCAAGAAATCTCTGGAATTAAGTTTATAGGTACTTGGGACCGCCGAAGAGGGCTGGGA 807  
Db 854 GAAATCAAGGAAGCTCTTGATTTATGTTTACAGGCAATTTGGAA---CGAAGAGGAATTTGGG 910  
Qy 808 TGGGCAAGATGCAATCTTCTGATGTCGATGACACTGCCATGGGTCTTTAGAAATCCTG 867  
Db 911 TGGGCGACACTAATCCATAGCAGATCTTGGAGACCACTGCTTTGGGATTTTCGATTTGCTT 970  
Qy 868 AGACTTCAATGATACAAATGATCTTTCAGATGTTCTGAGAGAAATTTTCAAGACGAGAAAGGA 927  
Db 971 CGGCTGCATAGGTACAAATGATCTCCAGCCATTTTGTGACAACTTTTCAAGATGCCAAATGGG 1030  
Qy 928 GACTTCTTTGCTTTCGCGTCAAAACGCAATTTGGTGTGACCGATATCTTAACTTTAT 987  
Db 1031 AAATTCATTTGCTCGACCGGCTCAATTCACAAAGATGTAGCAAGCATGCTGGAATCTTTAT 1090  
Qy 988 AGATGTTTCAAGTATGTTTTCGCGGAGAAAGATTAATGGAAGAGCTTAAGACCTTCACT 1047  
Db 1091 AGAGCTTCCAGCTCGCATTTCCGCGAGAAACATCTTCTGATGAAGCTTAAAGCTTCGCT 1150  
Qy 1048 ACAATGATCTCCAAATGCTTTTGGCCAAACAAACGCAATTTGATTAAGTGGGTGCTCAAG 1107  
Db 1151 ACTAAATATTTGAGAGAAAGCTCTTGAAGAAAGTGAGACTTCCAGTGTGATGGAACAACAAA 1210  
Qy 1108 AAGGATCTTCTCGAGAGGTGGATGCTATTAAGTATCCGTGGCATAGAGTATGCCA 1167  
Db 1211 CAAAACCTGAGCCAGAGATCAAAATGCGCGCTGAAGACTTCTTGGCATGCGAGTGTTCGG 1270  
Qy 1168 AGATTGGAGGCAAGAGTATACATAGAGCAATTTTGGATCAAAATGATGTCTGCTGGGGAAG 1227  
Db 1271 AGAGTGGAGCAAGAGATTAATGCTCAAGTGTATCGCCAGATTTATGACGCATAGCAAAA 1330

Search completed: August 25, 2004, 02:26:24  
Job time : 983.588 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 24, 2004, 06:59:45 ; Search time 5672.28 Seconds  
(without alignments)  
11797.935 Million cell updates/sec

Title: US-10-041-007-38

Perfect score: 2241

Sequence: 1 agtttcagtcacatggcgca.....ttttcgatcctgtggcgtag 2241

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estlin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	372.6	16.6	771	14	CF478561
2	307.6	13.7	730	13	BX682869
3	267.4	11.9	641	14	CF475347
4	262.4	11.7	2797	11	AY109636

5	231.4	10.3	476	10	BF517045
6	228.8	10.2	725	14	CF397293
7	214.8	9.6	583	13	BX681656
8	205.4	9.2	634	14	CF397946
9	204.4	9.1	725	14	CF470471
10	190.8	8.5	700	14	CF474786
11	190.4	8.5	745	14	CF477103
12	188.8	8.4	696	14	CF401316
13	186	8.3	581	13	BX677599
14	184.4	8.2	624	13	BU971902
15	184.2	8.2	688	14	CA223665
16	176.2	7.9	569	10	AW287754
17	175.8	7.8	601	9	AW043070
18	175.2	7.8	833	14	CB671661
19	174.8	7.8	638	14	CA203354
20	174.2	7.8	873	12	BG368010
21	172.8	7.7	669	14	CF479802
22	172.2	7.7	863	14	CB630278
23	171.8	7.7	706	14	CA917162
24	170.4	7.6	616	14	CF663845
25	169	7.5	576	9	AU299985
26	166.2	7.4	804	14	CF666338
27	165.2	7.4	506	12	BG317597
28	165.2	7.4	865	14	CB671672
29	164.4	7.3	697	13	BQ196773
30	164.4	7.3	740	14	CF477562
31	158.8	7.1	554	13	BU101395
32	157.4	7.0	533	13	BQ698297
33	156.6	7.0	599	13	BX680641
34	156.2	7.0	669	14	CF397861
35	156.2	7.0	682	10	CF397245
36	153.6	6.9	539	10	AW287756
37	146.8	6.6	591	14	CA305371
38	146.6	6.5	741	14	CF389118
39	146.4	6.5	840	14	CB624295
40	143	6.4	839	14	CB619396
41	143	6.4	844	14	CB619357
42	140.8	6.3	599	12	BG526917
43	140.4	6.3	385	10	AW697531
44	140.4	6.3	651	14	CF479884
45	135.4	6.0	843	14	CF208265

#### ALIGNMENTS

RESULT 1  
CF478561

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CF478561 771 bp mRNA linear EST 08-SEP-2003  
RTMW3\_20\_F04.g1\_A022 Well-watered loblolly pine roots WM3 Pinus  
taeda cDNA clone RTMW3\_20\_F04\_A022 5', mRNA sequence.

CF478561  
EST.  
Pinus taeda (loblolly pine)

Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Pinaceae; Pinus; Pinus.

1 (bases 1 to 771)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
Neale, D.

An EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of  
Forestry, University of Georgia; plant material prepared at the

CF208265 CAB20003\_

University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

```

FEATURES
    source
        1..771
            /organism="Pinus taeda"
            /mol_type="mRNA"
            /strain="CCLONES"
            /db_xref="taxon:3352"
            /clone="RTW3 20 F04 A022"
            /lab_host="DH10B-T1 Phage-resistant E. coli"
            /clone_lib="Well-watered loblolly pine roots W3"
            /notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 Mpa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."
```

ORIGIN

Query Match 16.6%; Score 372.6; DB 14; Length 771;  
Best Local Similarity 69.0%; Pred. No. 1.2e-97;  
Matches 510; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

Qy 1 ATGTTTCAGTCCGAGGCGGTAACGATCATCTGCATATGATACAGCTTGGGTG 60  
Db  
Qy 33 ATGTTTCAGAGGTATGGGTGATGGGGAACAATCCCTCGGTATGACACTGTTGGTA 52  
Db  
Qy 61 GCAAGAAATTCGGTCAATGACGGCTCTGTGTCACCCCAATTTCCCAACGCTTCAATGG 120  
Db  
Qy 93 GCAAGATTCAGCACTTGTATGGCTCTGACCACTCTCAGTTTCTCAACACTCAATGG 152  
Db  
Qy 121 ATTCTGAACATCAACTCCAGATGGCTCGTGGGTGAGGAGTGATTTTCGGCGTAT 180  
Db  
Qy 153 ATTCTTCAGATCAGTCAAGATGGCTCTGAGCACTCTCAGTTTCTCAACACTCAATGG 212  
Db  
Qy 181 GACAGAGTTTAAACACTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
Db  
Qy 213 GACAGAGTACTCACT 272  
Db  
Qy 241 ATTCAAGTGCAGAAAGGGGTTCAGTTTGTGAAACACATGTAAGAAATGAAGACGAA 300  
Db  
Qy 273 ATACAAGTGCAGAAAGGGTCTTGAATTTCTTCAAGAACACGCTGAAATGATGAGATGAA 332  
Db  
Qy 301 GCTGACATCAAGGCGCAAGTGGATTCGAGGTCTGTGTTTCTGCAATGTTAGATGAAGCA 360  
Db  
Qy 333 GCGGACATCAAGGCGCAAGTGGATTCGAGGTCTGTGTTTCTGCAATGTTAGATGAAGCA 392  
Db  
Qy 361 AAAAGCTTCGGATGTGATCTTCCTTATCACTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
Db  
Qy 393 AAACATTTATGTTGATCTGCTTATGATGTTGCTTATTAACCAATCAATTAAGAG 452  
Db  
Qy 421 CGCCAGAAAGCTTCAAGATTTCCCTCAATGTTTCTTCAATGTTTCTTCAATGTTTCTTCA 480  
Db  
Qy 453 GGGGAAGTAAAGCTTAAAGATTTCTCACTGATCTCTCTATACGGTTCACCAATATAC 512  
Db  
Qy 481 CTCTACTCTCTGAGGGTTTTCAGAGATGTGTGGACTGGCAGAGATCACAATCTTCAA 540  
Db  
Qy 513 TTGTATTTTGAAGGCTTCAAGAAATAGTGGAGTGGGCCCAAAATCATCAAACTTCAA 572  
Db  
Qy 541 TCAAGAGACGGATCAATTTTAACTCCCTCCGATCTACTCTGTTGTCTTCATGCACT 600  
Db  
Qy 573 TCCAAGGATGGATCGTTTCTCGGCTCGCCAGCATCTACAGCCGTGTATTCATCGGTACA 632  
Db  
Qy 601 CAAAACAAACGATGCCCTTCACTTTCTCAACTTCTGTCTCAGCAAAATTTGGCGACTAGCTT 660  
Db  
Qy 633 GGGNACCAAGTGTCTTGAATTTCTGAACTTGTGTTCTTAATGAAATTCGGAACCAAGCTG 692  
Db  
Qy 661 CCTTGCCATTAACCACTTGTATCTATTTTGAAGCGCTCTGGGCTGTGCGATACAGTTGAACGC 720

```

Db 693 CCGTCTCACTATCCGCTTAATCTATTGAACGCTGTGGCGCTTGATACTGTGAACGG 752
Qy 721 TTGGGAATCGATCGCTATT 739
Db 753 TTAGGATCGATCGGCATT 771

RESULT 2
BX682869 730 bp mRNA linear EST 06-NOV-2003
BX682869 Pinus pinaster differentiating xylem adult Pinus pinaster
cDNA clone 122H09 similar to Pinene synthase (EC 4.2.3.1), mRNA
sequence.
BX682869 1 GI:38196667
BX682869
EST.
Pinus pinaster
Pinus pinaster
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Pinaceae; Pinus; Pinus.
1 (bases 1 to 730)
Canton, F.R., Le Provoet, G., Garcia, V., Barre, A., Frigerio, J.-M.,
Paiva, J., Fereiro, P., Avila, C., Mouret, J.-F., Brach, J., de
Daruvur, A., Canovas, F.M. and Plomion, C.
Transcriptome analysis of wood formation in maritime pine
Unpublished (2003)
Contact: Frigerio JM
Genetique et Amelioration 69
INRA
route d'Arsachon 33612 Cestas CEDEX France
Email: Frigerio@pierrot.inra.fr
Email: Frigerio@pierrot.inra.fr.
Location/Qualifiers
1..730
/organism="Pinus pinaster"
/mol_type="mRNA"
/db_xref="ecotype: Corsican"
/db_xref="taxon:71647"
/clone="122H09"
/tissue_type="differentiating xylem"
/dev_stage="adult"
/clone_lib="Pinus pinaster differentiating xylem adult"
/notes="Vector: Uni-Zap XR lambda (Stratagene); Site 1: Eco
RI; Site 2: Xho I; A composite cDNA library was made with
mRNA isolated from normal, compression, opposite, early
and late wood of Maritime pine uni-directionally cloned
into Uni-ZAP XR using the ZAP-cDNA Synthesis kit
(Stratagene). pBluescript SK(-) plasmids were obtained by
in vivo mass excision. The nucleotide sequence of the
5' end was obtained by automated sequencing with the T3
primer by GENOME EXPRESS, Meylan, France"
```

ORIGIN

Query Match 13.7%; Score 307.6; DB 13; Length 730;  
Best Local Similarity 64.8%; Pred. No. 1.5e-78;  
Matches 473; Conservative 0; Mismatches 254; Indels 3; Gaps 1;

Qy 1331 TTGTAGCTGTGGAGAGATCGGATTCATATGATCTTACATTCACCGCCAGCGCGCTG 1390  
Db 1 TTGCGAGGTGGTGAATTCATCGGTTTCTCAGAGCTGGGCTTCACTCGGAACGTGTGG 60  
Qy 1391 TGGAAATGATTTCTCAGTGGCGGTATGATGTTGAGCCAGATTCGCTGCTGTAGAA 1450  
Db 61 CTGAATATATTTTTCACGAGGTCTATTTTGTGAGCCAGAGTTTCTGCTACCTGCGGAG 120  
Qy 1451 TTGCTATGCGAAGATCTTTCGCTCGAGTATTTCTAGACGATCTTTACGACACCCACG 1510  
Db 121 CCCTTTACACAAACCTTCTTAATTTCAAGTCTATATGATGATCTTATGACGCTCATG 180  
Qy 1511 GATCTCTGAGTATCTTAATTTCTCTGAAAGCGTCCGAAGATGGGATATCTCTGTC 1570  
Db 181 GAACTTTGGAATCTCAAGTTGTTTTCAGAACTCTGTGAAAGATGGGATCTATCTCTTGG 240

QY 1571 TGGATAGCGTTCCGGGATAATCAGTTGAAAGTTTCTCTAGGGCTGTACAAACACAGTGA 1630  
 Db 241 TAGATCAAAAT- -GCCACACAGCATGAAATATGTTTAAAGGGTTTTTACATACTTTTA 297  
 QY 1631 ATGGATTTGGAAAAGATGGACTCAAGGAACAAGCCGTGATGCTGGGGTATCTTCGAA 1690  
 Db 298 ATGAAATCGCTGAAGAGGGCGGAAGAGCGGGCGTGAATGCTAGGTTTACATTCAAA 357  
 QY 1691 AAGTATGGGAGGCTTGTCTGCAATCGTATACCAAGAGCCGAATGCTCGGCGACAAAGT 1750  
 Db 358 AAGTTTGGGAGGCTTGTCTGCAATCGTATACCAAGAGCCGAATGCTCGGCGAGTAGAT 417  
 QY 1751 ATGTGCGGACATTCACGAATATGTGGAATAATGTCGAAGAGTGTCTATAGCACTTGGGACAG 1810  
 Db 418 ACGTGCGGCTCTATGATGATATATAGGAACGCGAGTGTTCATAGCATTTGGGAACAG 477  
 QY 1811 TCGTACTTAACTCAATCTTTTTCACATGGAGAAATTTACTTCTGATTTACATTTTACAGCAAG 1870  
 Db 478 TGGTTCTCATCAGCGCTCTTTTTCACCGGGAGATTTCTACAGATGATATATCTTCCAAA 537  
 QY 1871 TAGACCTTCGGTCCAAATTTCTGCATCTTGTGTCTTTTGAAGTGGACGACTAATCAATGACA 1930  
 Db 538 TTGGTCCGGAATTCAGATTTCTATACCTCATGCGCTTTGACAGGGCGTCTGCTGACGACA 597  
 QY 1931 CCAAGACTTACAGGCGGAGAGAAACGGTGTGAATTTGGTTTCCAGGCTACAGTGTCTACA 1990  
 Db 598 CCAAAACTTATCAGGCTGAGAGAGGTCAGAGAGAGGTTCTTCTGCAAGTACAGTGTTTACA 557  
 QY 1991 TGAGGGAATAATCCGGAGTGCACAGAGGAAGCAAGCTCTTAAGTCAATGTTTATGTTATCATCG 2050  
 Db 658 TGAAGACCACTCTGAGATCTCGAAGAGAGAGCTCTCAACATGCTATATCTGTCATGG 717  
 QY 2051 ACAACGCAT 2060  
 Db 718 ATATGCTCT 727

## RESULT 3

CF475347  
 LOCUS  
 DEFINITION RTW2\_13 C12\_g1\_A021 Well-1 watered loblolly pine roots W2 Pinus taeda cDNA clone RTW2\_13\_C12\_A021 5', mRNA sequence.  
 ACCESSION CF475347  
 VERSION CF475347.1 GI:34492719  
 KEYWORDS EST.  
 SOURCE Pinus taeda (loblolly pine)  
 ORGANISM Pinus taeda  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.M., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J., Cannon, R., Owen, A. and Neale, D.  
 EST database from well-watered loblolly pine (Pinus taeda) roots  
 Unpublished (2003)  
 Other ESTs: RTW2\_13 C12\_b1\_A021  
 Contact: Cordonnier-Pratt MM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
 Seq primer: JENREV (GAGGAACAGCTATGACC).  
 Location/Qualifiers

## FEATURES

source 1..641  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /strain="CCLONES"  
 /db\_xref="taxon:3352"  
 /clone="RTW2\_13 C12\_A021"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots W2"  
 /notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 Mpa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."  
 ORIGIN

Query Match 11.9%; Score 267.4; DB 14; Length 641;  
 Best Local Similarity 67.9%; Pred. No. 9.3e-67;  
 Matches 389; Conservative 0; Mismatches 181; Indels 3; Gaps 1;  
 QY 1006 TTTCCGGGAGAAAGATAATGGAAGAGCTTAAGACCTTCACTACAATCATCTCCAAAAT 1065  
 Db 3 TTTCCGGGAGAAAGATCATGAGGGAAGCAAACTCTGTACGGAAAGGTATCTGAAGCAT 62  
 QY 1066 GCTCTTGGCCAAAACAAACGCAATTTGATAGTGGGCTGTCAAGAGGATCTTCTCTGGAGAG 1125  
 Db 63 GCTCTGGAGGATGTGATGCTTTGACAAATGGGCCATTTAAAAAGAAATCTTCGGGCGAGAG 122  
 QY 1126 GTGGAGTGTCTATAAGTATCGTGGCATAGAGTATGCCAAGATTTGGAGGCAAGAGT 1185  
 Db 123 GTGGAGTATTTCACTCAAAATATCCCTGGCATAGAGTATGCCAAGGCTGGAGGCCAGAGC 182  
 QY 1186 TACATAGAGCAATTTGGATCAAAATGATGTCTGGCTGGGGAAGACTGTGTATAAGATGCTA 1245  
 Db 183 TAGCTTGAATTTATGACCAAGGATTTCTGGATCGGAAATTTGTATTTGATGCCA 242  
 QY 1246 TATGTGAGCAACGAAAATAATTTGGAGCTGGCCAAATTTGACATTCATATATGTTGGCAGGCC 1305  
 Db 243 AATGTTTTCAACGAAAAATATTTGGAATTAGCAAAATTTGGATTTCAATCGACTGCAAGCT 302  
 QY 1306 TTACACAAAAGGAGACTCAACATGTCAGTGTGGGAGAGATTCGGGATTCATATGAT 1365  
 Db 303 TTACACAGACAGAGCTTCAGGATATTCGGAGGTGGTGGAAATTCATCGGGTTTCTCAGAG 362  
 QY 1366 CTTACATTCACCCCGCAGCGGCTGTG---GAAATGATTTCTCAGTGGCGGTTAGTATG 1422  
 Db 363 CTCGGCTTCACTCATCGGNAACGTGTGGCCGAATATATTTCTCAGCAGCGTCTCTTTTG 422  
 QY 1423 TTTGAGCCAGAAATTCGCTGTTAGAAATTCCTATGATGCGCAAGACTTCTTCCCTCGCAGTT 1482  
 Db 423 TTTGAACACAGAGTTTCTACCTGACGAGCGGTTTACACAAAAACTGCTATTTTTCACCGTC 482  
 QY 1483 ATTCTAGAGCTTTTACGACACCCAGGATCTCTCGATGATCTTAAATTTGTTCTCTGAA 1542  
 Db 483 CTATTAGATGATCTTTATGACAGCATGAACCTTTAGACAATCTTAACCTGTTTACAGAA 542  
 QY 1543 GCGGTCCGAGAGTGGGATATCTCTGTGCTGGAT 1575  
 Db 543 TCTGTGAAAAGATGGGATCTATCACTGGTAGAT 575

## RESULT 4

AV109636  
 LOCUS  
 DEFINITION Zea mays CL2355\_1 mRNA sequence.  
 ACCESSION AV109636  
 VERSION AV109636.1 GI:21213434  
 KEYWORDS HTC.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 2797)  
 Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whiteitt, M.S.,  
 Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.  
 Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
 Unpublished (2002)  
 2 (bases 1 to 2797)  
 Coe, E.H.  
 Direct Submission  
 Submitted (25-APR-2002) Maize Mapping Project, University of  
 Missouri, Columbia, MO 65211, USA  
 If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB and may be found by BLAST  
 searching at MSU, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR,  
 www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
 www.zmdb.iastate.edu.

FEATURES  
 Location/Qualifiers  
 1..2797  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="MaizeDB:630936"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 configs to seed Dupont contigs; this resource was  
 assembled by Dupont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

ORIGIN  
 Query Match 11.7%; Score 262.4; DB 11; Length 2797;  
 Best Local Similarity 45.5%; Pred. No. 6.3e-65;  
 Matches 674; Conservative 0; Mismatches 78; Indels 21; Gaps 6;

1 ATGTTTTCAGTCCATGGGCGAGGTGAACGAATCCATCTGCATATGATACAGCTTGGGTG 60  
 402 ATGCTACGGTTCATGAACGACGGGGATACCGGCTCGGCTAGCAGCGCGTGGGTG 461  
 61 GCAAGAAATTCGGTCAATGACGGCTCTGGTGACCCCA-----TTTCCCAACGCTT 114  
 462 GCGATGGTCCGNN 521  
 115 CAATGGATTCGAACAATCACTGCCAGATGCTGCTGGGGTGAGGAGTGCAATTTTCTG 174  
 522 CGCTGGATCGTGACCAACGCTGCCGACGCTCTGGGGGACTCGGCCCTGTTCTCC 581  
 175 GCGTATGACAGAGTTTAAACACTCTGCGCTCGCTCTCACTCAATATGGAATAG 234  
 582 GCTACGACCGCATGATCAACCCCTCGCTGCTGCTGCTGACCAAGTGGTGGCTG 641  
 235 GCGGCAATTCAGTGACGAAGAGGGTTGATGTTGTGAGAAACACATGAGAAATAG 294  
 642 GAGCCCGGAGTGGCAGCGGGGCTCTCGTTCCTGCAGAGACATGAGGCTAGCG 701  
 295 GACGAAGCTGACAAATCACAGGCGCAAGTGGATTCGAGTCTGTTTCTGCAATGTTAG 354  
 702 GAGGAGGCGGAGTCCGATGCCCATCGGCTTCGAGATCGCTTCCTCTCATCCAG 761  
 355 GAAGCAAAAGCTTGGGATTG---GATCTCTCTTATCACCTCCCTTTCATCTCCAAATC 411  
 762 ACGGCTAGGACCTGGGCGCTGCTGACTTCCTCGGTGCGGACACCGCGGCTGCAGACNTA 821  
 412 CACCAAAAGCGCAGAAAAGCTTCAAAAGATTCCCTCAATGTTCTTCAATACCATCAG 471  
 822 TAGCCCAACAGGAGNAGTCAGCTGACGGGATCCAGAGGACATGATGACAGGGTCCCG 881  
 472 ACGCGGTGCTTACTCTCTGAGGGGTTTGAAGATGTTGAGTGTGGTGGACTGGCAAGATCACA 931

Db 882 ACGTCCATCCTGCACAGNN 941  
 QY 532 AATCTTCAATCAAGAGAGCGGATCATTTTAAAGCTCCCTGCATCTACTCTGTGTCTTC 591  
 Db 942 NNNNNNNCT---GCGACNN 998  
 QY 592 ATGCACACTCAAAAACAAACGATGCTCCACTTTTCTCACTTCTGCTCAGCAAAATTTGGC 651  
 Db 999 ATGCAAAACCGGTGNN 1058  
 QY 652 GACTAGTTCCTTGGCATTACCCACTTGTATCTATTGAAAGCGCTCTGGCTGCGGATACA 711  
 Db 1059 NNNNNAGTCCCAATGTTTATCGGTGATCTTTTCGAGACATCTGGGTGTGGATCGG 1118  
 QY 712 GTTGAACGCTTGGGATCGATCGTATTTCAGAAAAGAAATCAAGAATCTCTGGATTAC 771  
 Db 1119 TTTGAGCGACTGGGATCTCCCGCTACTTCTCAACGAGAGATTGAGCAGTGCATGACTAT 1178  
 QY 772 GTTTATAGGTACTTGGGACCGCGAAGAGCGGTGGGATGGGCAAGATGCAATCTCTATCT 831  
 Db 1179 GTGAACAGGACACTGG---ACTGAAGATGGGATTTGCTGGGCTAGGNNNNNNNNNN 1235  
 QY 832 GATGTCGATGACACTGCCATGGTCTTAGAATCTCTGAGACTTCTCATGATACAAATGATCT 891  
 Db 1236 NNN 1295  
 QY 892 TCAGATGCTTCTGGAGAAATTTTCAGAGACGAGAAAGGAGACTTCTTTTCTGCTTGGCGG 951  
 Db 1296 CCAAGTGTGTTTAAAGACTTT---GAGAAAGATGGAGAGTCTTTTCTGTTGGGCGCA 1352  
 QY 952 ACGCAAAATTTGGTGACCGGATTAATCTTAACTTTATAGATGTTTCACAGATGATGTTTCCG 1011  
 Db 1353 TCGACTCAAGCGGCTCACTGGGATGTATACTTCAACAGAGGCTCTCAGATGAAGTTTTC 1412  
 QY 1012 GCGAAAAGATAATGGAAGAGCTTAAGACCTTCTACTACAAATCTATCTCCAAAATGCTCTT 1071  
 Db 1413 GGAGAGATGATTTGATCTGCTGCTAGGGTTTCTCGTATGATGTTTNNNNNNNNNNNNNN 1472  
 QY 1072 GCCAAAACCAACGATTTGATTAAGTGGGCTGCTCAAGAGGATCTTCTGGAGAGGTGGAG 1131  
 Db 1473 NNNNNNNCATGATCCGCTGATAAATGGATCTGTTCCCAAGGATCTACCTGGCGAGGTGCA 1532  
 QY 1132 TATGCTATAAGTATCCGTGGCATAGAGTATGCCAAGATTTGGAGGCAAGAGTTTACATA 1191  
 Db 1533 TATACACTAGACTTCCCTTGGTATGCAAGCTTGGCTCTGCTAGAGGCAAGACCTATCTA 1592  
 QY 1192 GAGCAAT---TTGGATCAAAATGATGTTCTGGTGGGGAAGACTGTGTATAAGATGCTATAT 1248  
 Db 1593 GATCAATATGGTGGTAAAGATGACGTTTGGATTGGAAGAGACACTCTACAGGATGCTCTT 1652  
 QY 1249 GTGAGCAACGAAAATATTTGGAGCTGGCCAAATTTGGACTTCAATATGTTGAGGCTTTA 1308  
 Db 1653 GTGAATAACGNN 1712  
 QY 1309 CACCAAAAGGAGACTCAACACATTTGCTGAGTGGTGGAGAGAAATCGGATTCATGATCTT 1368  
 Db 1713 CATCAGCTTGAATGTTAATGGGCTGCAAGCTGTACAGGATAAATGCTTCCGCTTTCAG 1772  
 QY 1369 ACATTCACCGCCAGCGGCTGTGGAATGTTTCTCAGTGGCGTTAGTATGTTTGGAG 1428  
 Db 1773 GGAGTAGAACCAACAGATGTTTTTAAGATCTTACNNNNNAGTCTGCTTGCANNNGAA 1832  
 QY 1429 CCAGAATTCGCTGTTGTAGAAATTTGCTATGCAAGACTTTC 1469  
 Db 1833 CTTAGCCGTGCTGCTGAGCGGCTTGATGGGCTAGAACGTC 1873

RESULT 5  
 BP517045  
 LOCUS  
 DEFINITION

BP517045 476 bp mRNA linear EST 07-MAY-2003  
 NXSI\_008\_E08\_F NXSI (Nsf Xylem side wood Inclined) Pinus taeda cDNA  
 clone NXSI\_008\_E08\_5, similar to Arabidopsis thaliana sequence  
 At4g02780 ent-kaurene synthetase A - like protein see

http://mips.gsf.de/proj/thal/db/index.html, mRNA sequence.  
BF517045.1 GI:11604507  
EST.  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Pinus taeda (loblolly pine)  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
Sederoff, R.  
1 (bases 1 to 476)  
Molecular Basis of Wood Formation in the Pine Megagenome  
Unpublished (2000)  
Contact: Sederoff, Ron  
Forest Biotechnology  
North Carolina State University  
840 Main Campus Drive, Centennial Campus, Campus Box 7247, Raleigh,  
NC 27695, USA  
Tel: 919 515 7800  
Fax: 919 515 7801  
Email: ron.sederoff@ncsu.edu, jerri.johnson@ncsu.edu  
Please see http://web.ahc.umn.edu/biodata/nsfpine/ for further  
information.  
Seq primer: T3.

#### FEATURES

source  
1..476  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="Coastal plain loblolly pine from North Carolina"  
/db\_xref="taxon:3352"  
/clone="NXSI\_008 E08"  
/tissue\_type="Xylem"  
/cell\_type="Side"  
/dev\_stage="Juvenile"  
/lab\_host="X11-Blue"  
/clone\_lib="NXSI (Nsf Xylem Side wood Inclined)"  
/notes="Vector: Bluescript SK; Site 1: Eco RI; Site 2:  
XhoI; The library is from early (spring) wood, taken from  
three six-year old trees (three different genotypes), in  
the juvenile phase. These trees were induced to form side  
wood by bending to a 45 degree angle and tying them to the  
ground. Differentiating xylem was harvested from the sides  
of the inclined stems, and a mixture of all three  
genotypes was used for the library. oligo-dT primed cDNA  
was directionally cloned into the EcoRI-XhoI Bluescript SK  
vector arms. NOTE: The sequences contain a 'cDNA adapter'  
between the EcoRI site and the start of the EST. The  
adapter sequence is 'AATTGGCAGCAG'."

#### ORIGIN

Query Match 10.3%; Score 231.4; DB 10; Length 476;  
Best Local Similarity 66.4%; Pred. No. 3.1e-56;  
Matches 316; Conservative 0; Mismatches 160; Indels 0; Gaps 0;  
QY 240 CATTCAAGTCACAAAGGGGTTGAGTTTGAGAAACACATCGAAGAATGAAGCAG 299  
DB 1 CACACAAGTACAAAGGGGATTGAATTTTTCAGACACAGCTGGGAAAATAGAGAGGA 50  
QY 300 AGCTGCAATCACAGGCCAGTCGATTCGAGGTCGTGTTTCTCGCAATGTTAGATGAAGC 359  
DB 61 GGCTGATGCCACAGGCCAAGTGGATTGAAATAGTATTTCCTGCAATGCTAAAGAGC 120  
QY 360 AAAAAGCTGGGATGGATCTTCCTATATCACTCCCTCTTCATCTCCCAATCCACCAAAA 419  
DB 121 TAAAGCCTTAGGCTGGCTCTGCTTATGAATTCCTTCATCCAGCAATCATTTGAAA 180  
QY 420 GCGCCGAGAAAGCTTCAAAGATTCCTTCATGTTCTTCATACCCNTCAGAGCGGTT 479  
DB 181 GCGGAGGCTAAGCTTCAAGGGCTTCCTCTGTATCTTCTATGCTTTACCAACAGTT 240  
QY 480 GCTCTACTCTCGAGGGTTTGCAAGATGTGGTGGACTGGCAGAGATCACAATCTTCA 339  
DB 241 ACTGTATTCTTGGAGNCCCTGCAAGAAATAGTAGACTGGGAAATATATGAATTTCA 300

QY 540 ATCAAGACGAGGATCATTTTAAAGTCCCTGCTGCTACTGCTTGTGTCTTCATGCAC 599  
DB 301 ATCCANNNTGGGTCAATCTCTAGCTGGCAGCATCTACAGCGGCTGTGTTCATGCGTAC 360  
QY 600 TCAAAACAAACAGATGCTCCACTTTCTCAACTTCGCTCAGCAAAATTTGGCGACTACGT 659  
DB 361 AGGACACAAAATAGCTGNGNNNTCTTGAATTTTCTTGAAGAAATTCGGAACCATGT 420  
QY 660 TCCTTCCCAATACCCACTTGATCTATTGAACGCCCTCTGGGCTGTGATACAGTTG 715  
DB 421 TCCTTGTCTACTATCCGCTTGANNCTTTGAACGNNNGNNGCGTGTGATAGTCTGNG 476

#### RESULT 6

CF397293  
LOCUS  
DEFINITION  
RTDS3\_2\_A09\_g1\_A022 Drought-stressed loblolly pine roots DS3 Pinus  
taeda\_cDNA clone RTDS3\_2\_A09\_A022 5', mRNA sequence.  
CF397293  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
Pinus taeda (loblolly pine)  
ORGANISM  
Pinus taeda  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.  
REFERENCE  
AUTHORS  
1 (bases 1 to 725)  
Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
Neale, D.  
An EST database from drought-stressed loblolly pine (Pinus taeda)  
roots  
Unpublished (2003)  
Other ESTs: RTDS3\_2\_A09\_b2\_A022  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmp Pratt@uga.edu  
RNA prepared and library constructed by W. Walter Lorenz, School of  
Forestry, University of Georgia; plant material prepared at the  
University of Florida; sequencing done in the Laboratory for  
Genomics and Bioinformatics, University of Georgia. Sequence ends  
have been trimmed to exclude vector and regions below Phred quality  
16. Three-prime sequences are presented as their reverse complement  
and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

#### FEATURES

source  
1..725  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCLONES"  
/db\_xref="taxon:3352"  
/clone="RTDS3\_2\_A09\_A022"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The  
library was prepared from polyA+ RNA from drought-stressed  
loblolly pine (Pinus taeda) roots. Water was withheld from  
ramet clones until predawn needle water potential reached  
-1.75 MPa. On day 7 roots were harvested for RNA  
isolation. Double-stranded cDNA was cloned  
unidirectionally into pSL1180. Inserts excised with EcoRI  
(5' end) and XhoI (3' end)."

#### ORIGIN

Query Match 10.2%; Score 228.8; DB 14; Length 725;  
Best Local Similarity 58.3%; Pred. No. 2.2e-55;  
Matches 420; Conservative 0; Mismatches 297; Indels 3; Gaps 1;  
QY 1130 AGTATGCTATAAGTATCCGTGCGTAGAAGTATGCCAAGATTGGAGGACAGATTACA 1189

Thu Aug 26 08:27:54 2004

```

/mol_type="mRNA"
/db_xref="taxon:71647"
/clone="RS62E09"
/tissue_type="root"
/dev_stage="6 weeks old seedling"
/lab_host="SOLR"
/clone_lib="RS"
/notes="Vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A three weeks drought stress treatment was applied by lowering the osmotic potential of the nutrient solution to -0.45 MPa using 170 g/l of polyethylene glycol as an osmoticum. A mixture of genotypes were used. Oligo-dr primed cDNA was used. directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBluescript phagemid"

```

ORIGIN

```

Query Match          9.6%; Score 214.8; DB 13; Length 583;
Best Local Similarity 63.6%; Pred. No. 2.6e-51;
Matches 361; Conservative

QY 1292 ATATGTCAGGCGCTTACACCAAAAGGAGACTCAACACATTTGTCAGCTGGTGAGAGAAT 1351
Db 1 ATCGTGTGCAATTTTACACAGACAGAGCTTCAAGATATCGGAGG--TGGTGAATCAT 58

QY 1352 CGGGATTCAATGATCTTACATTCACCCGCGCGCTGTGGAAATGATTTCTCAGTGG 1411
Db 59 CGGGTTTCTCAGAGCTGGGCTTCACTCGGGAACGTGTGGCTGAAATATATTTTTCACAG 118

QY 1412 CGGTAGTATGTTTACGCGAGATTCGCTGCTGTAGAAATTCCTATGCCAAGACTTCTT 1471
Db 119 CGTCATTTTGTGTGAGCCAGAGTTCCCTACCTCCGAGCGCTTTACACAAAACCTTCTA 178

QY 1472 GCCTCGCAGTTATTTCTAGAGCATCTTTACGACACCCACCGAGTCTCTGGATGATCTTAAT 1531
Db 179 ATTTACGGTCATATTAGATGATCTTTATGACGCTCATGGAATTTGGACAATCTCAAGT 238

QY 1532 TGTCTCTGAAGCGGTGGAAGATGGGATATCTCTGCTGGATAGCTTCGGGATAATC 1591
Db 239 TGTCTTCAAGATCTGTGAAAGATGGGATCTATCTTGGTAGATCAATGCCACAA--G 295

QY 1592 AGTTGAAGTTTGTCTCTAGCGGTGTACAAACACAGTGAATGGATTTGGAAGATGGAC 1651
Db 296 ACATGAAATATGTTTTTAAGGGTTTACATATCTTTTAAATGAAATCGCTGAAGAGGCG 355

QY 1652 TCAAGGAAACAGCGCGTGTATGCTGGGCTATCTTCGAAAGATATGGAGGCTTGTCTCG 1711
Db 356 GGNAAAGGCGAGCGGCGTGTATGCTAGTTTACATTTCAAAAGTTTGGAGGTCAGCTAG 415

QY 1712 CATCTATACCAACCAAGCCGAATGCTCGGACGAAAGTATGTGCGCAATTTCAACGAAT 1771
Db 416 GAGCGTATACCAAGAGAGAGATGCTGTCAGCTAGATACGCTGCGCTCTATGATGAAT 475

QY 1772 ATGTGGAATAATGCCAAAGTGTCCATAGCACTTGGCAGCTGCTACTTAATCAATCTTT 1831
Db 476 ATATAGGGAACCGAGTGTTCATATAGCAATTTGGGAACAGTGTCTCATCAGCGCTTTT 535

QY 1832 TCAGTGGAGAATTACTTCTCTGATTACAT 1859
Db 536 TCACCGGGGAGATTTCTTACAGATGATAT 563

```

RESULT 8

```

CF397946
LOCUS
DEFINITION
CF397946
ACCESSION
CF397946
VERSION
CF397946.1
KEYWORDS
EST.
SOURCE
Pinus taeda (loblolly pine)
ORGANISM
Pinus taeda

```

CF397946 634 bp mRNA linear EST 29-AUG-2003  
RTDS3 23 G06.g1\_A022 Drought-stressed loblolly pine roots DS3 Pinus  
taeda cDNA clone RTDS3\_23\_G06\_A022 5', mRNA sequence.

```

Db 2 AATAGCCCTGAATATCTTCTGGCATGCCAGTGTCCAGAGTGGAAAGCAAAAGATACT 61
QY 1190 TAGAGCAATTTGGATCAAAATGATGCTGCTGGGAGAGCTGTGTATAGATGCTATATG 1249
Db 62 GTCAGTGTATAATCCGATATGATGATGTTTACGAAAAAGTGTGTACAGCTGTCCAG 121
QY 1250 TAGACCAACAAAATATTTGGAGCTGGCCAAATGGACTTCAATATGTCGAGCGCTTAC 1309
Db 122 TGAACACACAAAAGCTCTCGAAATTTGGCAAAATAGATTTCAACATTTCCAGTCCATCC 181
QY 1310 ACCAAAGAGAGACTCAACACATTTGTCAGCTGTGGAGAGATTCGGAATTCATGATCTTA 1369
Db 182 ACCAAGAGAGATGAAGAGTGTACCAGCTGTGTTAAAGATTCGGGTTGCCCTCTATCA 241
QY 1370 CATTCACCCGCGAGCGGCTGTGGAAATGATTTCTCAGTGGCGGTAGTATGTTTGAGC 1429
Db 242 CTTTTCAGCGGAGACCGCTGGAATTTCTCTTGTATAGCGCTGTGTACGATGAGC 301
QY 1430 CAGAAATTCGCTGTTAGAAATTCCTATGCCAGAGCTTCTTCCCTCGCAGTTATCTTAG 1489
Db 302 CCCAGTATGCCAATGCAGATTTCTCTTACAAAAGTGGCGTCTTGCAGACTGTTCTGG 361
QY 1490 ACGATCTTTACGACCCAGATCTCTGATGATCTTAATTTCTCTGAAAGCGGTCC 1549
Db 362 ACGATATGATGACACTTATGGGACCTCGATGAATTTGAAGCTGTTTCACTGAGCTGTCA 421
QY 1550 GAGATCGGATATCTCTGTGCTGATAGCGTTCCGGGATATCAATGTTGAAAGTTTGCTTC 1609
Db 422 GAAATGGGACCTCTCTTATACGGAACCTTCCAGACTAT--ATGAAGCTATGTTTACA 478
QY 1610 TAGGGCTGTACAAACAGTGAATGGATTTGGAAGATGGACTCAAGGACCAAGCGCGTG 1669
Db 479 AAATCTATTATGACATAGTTTCAAGAGTGGCGTGGGAGGAGAGAAAGGAGACGTG 538
QY 1670 ATGTGCTGGCTATCTTTCGAAAGTATGGAGGCTTCTGCTGATCGTATACCAAGAG 1729
Db 539 AATTAGTTAGCTTTTTCGAAAGGAGTGGAGGATATCTTCTGGGTTATTTATGAAGAAG 598
QY 1730 CCGAATGCTCGGCAAGATATGTCCGACATTCACAGATATGTGGAATGCGCAAG 1789
Db 599 CTGAATGCTAGCAGCTGATGTGCCAGCTTGGACGAATACATAAAGATGGAATCA 658
QY 1790 TGTCCATAGCACTTCGGACAGCTGCTACTAACTCAATCTTTTCTACTGGAGAAATTTCTC 1849
Db 659 CTTCCATTGGCCACGTATACTTCTTTTGAAGTGGAGTATGATATGGAGGAGCACTCC 718

```

RESULT 7

```

BX681656
LOCUS
DEFINITION
BX681656
ACCESSION
BX681656
VERSION
BX681656.1
KEYWORDS
EST.
SOURCE
Pinus pinaster
ORGANISM
Pinus pinaster
REFERENCE
1 (bases 1 to 583)
AUTHORS
Frigerio, J. and Plomion, C.
TITLE
Identification of water-deficit responsive genes in Maritime pine
(Pinus pinaster Ait.) using an EST approach
JOURNAL
Unpublished (2002)
COMMENT
Contact: Frigerio JM
Genetique et Amelioration 69
INRA
route d'Arcachon 33612 Cestas CEDEX France
Email: Frigerio@pierrotton.inra.fr
Email: Frigerio@pierrotton.inra.fr
Seq primer: T3
Location/Qualifiers
1..583
/organism="Pinus pinaster"

```

FEATURES

source

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus. 1 (bases 1 to 634)

Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.M., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.

An EST database from drought-stressed loblolly pine (Pinus taeda) roots

Unpublished (2003)

Other ESTs: RTDS3\_23\_G06\_b1\_A022

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAACAGCTATGACC).

Location/Qualifiers

1. .634

/organism="Pinus taeda"

/mol\_type="mRNA"

/strain="CCLONES"

/db\_xref="taxon:3352"

/clone="RTDS3\_23\_G06\_A022"

/lab\_host="DH10B-T1 phage-resistant E. coli"

/clone\_lib="Drought-stressed loblolly pine roots DS3"

/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from drought-stressed loblolly pine (Pinus taeda) roots. Water was withheld from ramet clones until predawn needle water potential reached -1.75 MPa. On day 7 roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 9.2%; Score 205.4; DB 14; Length 634;

Best Local Similarity 60.4%; Pred. No. 1.6e-48;

Matches 357; Conservative 0; Mismatches 231; Indels 3; Gaps 1;

QY 1208 ATGATCTCGCTGGGAGACGTGCTATAGATGCTATATGTCAGCAGCAAGAAATATT 1267

DB 9 ATTATGCATGTTTACAAAAGTGTTTACAAGCTGTCCACGCTGAAACGAAAGCTCC 58

QY 1268 TGGAGCTGCGCAATTTGGACTTCAATATGTCAGGCTTTACACCAAAAGGAGACTCAAC 1327

DB 69 TGAATTGGCAAAATTAGATTTCACATTTGTCAGTTCATCCACCAAGAGAGATGAGA 128

QY 1328 ACATTGTCAGCTGGTGAGAGATCGGAGATTCATATGCTTCAATTCATCCCGCCAGCGGC 1387

DB 129 GTGTTACCAAGCTGTTTAAAGATTGGGGTGGCTCTTCTACCTTTTCACGGGAGAGAC 188

QY 1388 CTGTGGAATGATTTCTCAGTCGGCGGTAGTAGTTGTTGACGAGAAATTCGCTGCTGTA 1447

DB 189 CGCTGGAATTTACTTCTTTGATAGCGCTGTGATGAGCCCGCCAGTATGCCAAATGCA 248

QY 1448 GAATTCCTATGCGAGAGCTTCTTGCTCGCAGTTATCTTAGACGATCTTTAGCACCC 1507

DB 249 GATTTCTTTTACAAAGTGGCGTGTGTCAGACTGTCTTGGACGATATGTATGACACTT 308

QY 1508 ACCGATCTCTGGATGATCTTTAAATTTGTTCTCTGAAGCGGTCCGAAGATGGGATCTCTG 1567

DB 309 ATGGGACCTCGATGATTAAGACTGTTCAGTGGCTGTGACAGATGGGACTCTCTT 368

QY 1568 TGCTGATAGCGGTTCGGGATAATCAGTTGAAAGTTTGTCTTAGGGCTGTACAAACAG 1627

Db 369 ATACGAAAACCTTCCAGACTATA---TGAAGCTATGTTACAAAATCTATTATGACATAG 425

QY 1628 TGAATCGATTTGGAAAAGATGGACTCAAGGAAACAAGCCCGTGTGCTGGGTATCTTC 1687

Db 426 TTCACGAAGTGGCTGGGAGGACAGAGGAACAGGACGTGAATTACTTATGCTTTTCA 485

QY 1688 GAAAAGTATGGAGGGCTTGTGCGATCGTATACCAAGAGCGGAATGTGCGCAGCAA 1747

Db 486 GAAAGGATGGGAGGATTAATCTCTGGTTATTATGAAGAAGCTGAATGGCTAGCAGTG 545

QY 1748 AGTATGTCGCGACATTCACGAATATGTGAAAATGCCAAAGTGTCCATAG 1798

Db 546 AGTATGTCGCCAGCTTGGAGATACATATAAGAATGGAATCACTTCCATTG 596

RESULT 9

LOCUS CF470471

DEFINITION RTDS1\_17\_C03\_g1\_A015 Drought-stressed loblolly pine roots DS1 Pinus taeda cDNA clone RTDS1\_17\_C03\_A015 5', mRNA sequence.

ACCESSION CF470471

VERSION CF470471.1 GI:34487843

KEYWORDS EST.

SOURCE Pinus taeda (loblolly pine)

ORGANISM Pinus taeda

REFERENCE 1 (bases 1 to 725)

AUTHORS Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.M., Dean, J., Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and Neale, D.

TITLE An EST database from drought-stressed loblolly pine (Pinus taeda) roots

JOURNAL Unpublished (2003)

COMMENT Other ESTs: RTDS1\_17\_C03\_b1\_A015

Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology

Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA

Tel: 706 542 1860

Fax: 706 583 0210

Email: mmpratt@uga.edu

RNA prepared and library constructed by W. Walter Lorenz, School of Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: JENREV (CAGGAACAGCTATGACC).

Location/Qualifiers

1. .725

/organism="Pinus taeda"

/mol\_type="mRNA"

/strain="CCLONES"

/db\_xref="taxon:3352"

/clone="RTDS1\_17\_C03\_A015"

/lab\_host="DH10B-T1 phage-resistant E. coli"

/clone\_lib="Drought-stressed loblolly pine roots DS1"

/notes="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from drought-stressed loblolly pine (Pinus taeda) roots. Water was withheld from ramet clones until predawn needle water potential reached -1.75 MPa. On day 7 roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 9.1%; Score 204.4; DB 14; Length 725;

Best Local Similarity 58.9%; Pred. No. 3.4e-48;

Matches 371; Conservative 0; Mismatches 256; Indels 3; Gaps 1;



Forestry, University of Georgia; plant material prepared at the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.  
Seq primer: JENREV (CAGGAACAGCTATGACC).

Location/Qualifiers

# FEATURES

source

1..700  
/organism="Pinus taeda"  
/mol\_type="mRNA"  
/strain="CCIONES"  
/db\_xref="taxon:3352"  
/clone="RTW2\_7\_B11\_A021"  
/lab\_host="DH10B-T1 phage-resistant E. coli"  
/note="Vector: pSL1180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/-0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."

# ORIGIN

Query Match 8.5%; Score 190.8; DB 14; Length 700;  
Best Local Similarity 55.6%; Pred. No. 3.3e-44;  
Matches 388; Conservative 0; Mismatches 307; Indels 3; Gaps 1;

QY 1384 CGGCTGTGGAATGATTTCTCAGTCGCGGTAGTATGTTTGAGCCAGAAATTCGCTGCT 1443  
DB 2 CGTCACGTGGAATACTACACTTTAGCAGCTTGCAATGATCCTTAACATCTCTGGC 61  
QY 1444 TGTAAGATTTGCTATGCGCAAGACTTCTTGCTCGCAGTTATTTAGAGCACTTTACGAC 1503  
DB 62 TTTTCGACTAGGATTTGGTAAATAAGTCATATGATCAGGATTTCTGACGATATCTACGAC 121  
QY 1504 ACCACGATCTTGAGTATCTTAAATTTCTCTGAAGCGGTCGGAAGATGGGATATC 1563  
DB 122 ACCTTCGGAACAATGGAGGAGCTCGAATCTTTAACCGCAGCGTTTAAGAGATGGGATCG 181  
QY 1564 TCTGTGCTGATAGCGTTTCGGGATTAATCAGTTGAAAGTTTGTCTTAGGGCTGTACAC 1623  
DB 182 TCTTCGATAGAGTGTCTCCAGATTAT--ATGAAGGAGTGTATGCGGTTTACGAC 238  
QY 1624 ACAGTGAATGATTTGGAAAGATGCGATCAAGGACAGGCGGTGATGCTGGGCTAT 1683  
DB 239 AACATCAACGAAATGSCACGAGAGCGCGAGAAAATTCAGGGCTGGATACAGTCAGCTAT 298  
QY 1684 CTTCCGAAAGTATGGAGGCTTGTCTCGCATCGTATATACCAAGAGCCGAAATGGTCGCA 1743  
DB 299 GCTCGAAATCTTTGGAGGCTTTTATGCTTATATACAGAGCCAAAGTGGATTTC 358  
QY 1744 GCAAAAGTATGTCGCGACATTCACGAATATGGAATAATGCCAAAGTGTCCATAGCACTT 1803  
DB 359 AGTGGTTATCTCCACAGTTCGACGAGTACCTCGAGAATGGGAAGGTGAGCTTCGGCTCT 418  
QY 1804 GCGACGTGCTACTAACTCAATCTTTTCTGAGGAAATTAATCTCTGATTAATTTTA 1863  
DB 419 CGCATAAACACGCTCGAACCCATGCTGAGCTTTGGGGTTTCTCTTCGCGCTCGAATCTG 478  
QY 1864 CAGCAAGTATAGCTTCGCTCCAAATTTCTGCAATTTGTGCTTTGATGAGACGACTATC 1923  
DB 479 CAGGAATTTGACTTTCCATCGAAATTCATGATTTGATATGTGCCATCTCTTCGACTGAA 538  
QY 1924 AATGACCAAGACTTACAGGCGGAGAGAAACCGGTGTAATTTGTTTCCAGCGGTACAG 1983  
DB 539 GGTGACACTCAATGCTACAGGCTGACAGGCGCGTGGAGAGAGCTTGGCGGTATCG 598  
QY 1984 TGCTACATGAGGGAATAATCCGAGTGCACAGAGAGAAAGCTCTTAAGTCATGTTTATGGT 2043  
DB 599 TGTATATGAAGACCATCTCTGGAATAACAGAGAGAGATGCTGTCAATCAAGTCAATGCT 658

QY 1220 TGGGGAAGACTGTGTATAAGATGCTATATGTGACCAACGAAATAATTTGGAGCTGGCCA 1279  
DB 5 TAGCAAAAAGTGTTTACAGCTGCCCTACGTGAACCAACGAAAGCTTCTAGAGTTGGCAA 64  
QY 1280 AATTGGACTTCAATATGGTGGAGGCTTACACAAAGAGAGTCAACATTTGTCAGCT 1339  
DB 65 AATTAGATTCAACATTTGTCAGTCCATCCACCAAGAGAGATGAATAATGTTTACCAGCT 124  
QY 1340 GGTGGAGAGATCGGATTCATGATCTTACATTCACCCGCCAGCGCTGTGGAATGT 1399  
DB 125 GGTTTAAAGATTGCGGGTTCCTTATTCACCTTTGACCGGAGAGCGCTGGAATCT 184  
QY 1400 ATTTCTCAGTGGGGTTAGTATGTTGAGCCAGATTCGCTGCTGTAGAAATTCCTCTATG 1459  
DB 185 ACTTCTGTAGCGCTGGTACGTATGAGCCAGTATGCGCAATGCGAATTCCTCTTTA 244  
QY 1460 CCAAGACTTCTGCTCGCAGTTATTTCTAGACGATCTTTACGACACCCAGGATCTCTGG 1519  
DB 245 CAAAAGTGGCGTGTTCGCTGACTGTTCTGGACGATATGATGACACTTATGGGACCTTG 304  
QY 1520 ATGATCTTAAATGTTCTCTGAAGCGTCCGAAGATGGGATATCTTGTGCTGGATAGCG 1579  
DB 305 ATGAATTTGAAGCTGTTTACTGAGCGCTGAGAAGATGGGACCTCTCTTATACAGAAAACC 364  
QY 1580 TTCCGGATAATCAGTTGAAAGTTTGTCTTCTAGGCGGTGACACACAGTGAATGATTTG 1639  
DB 365 TTCCAGACTAT--ATGAACATATGTTACAAAATCTATTATGACATAGTTCACGAATGG 421  
QY 1640 GAAAGATGGAATCAAGGACCAAGCGGTGATGCTGGGCTATCTTCGAAAAGTATGGG 1699  
DB 422 CCTGGAGGCGAGAGAGGACAGGGCGGTGAATTTGGTTAGCTTTTCAGAAAGGATGGG 481  
QY 1700 AGGCTGTGCTCGCATCGTATACCAAGAGCGGAATGCTGGCAGCAAGTATGTCGCA 1759  
DB 482 AGGATTATATTTCTGAGTTATTATGAAGAAGCTGAATGTTAGCACTGAGTATGTGCCCA 541  
QY 1760 CATTCAGCAATATGTGAATAATCCAAAGTGTCCATAGCACTTGCAGACAGTCTGACTAA 1819  
DB 542 GCTTGAACGATACATCAAGAACGGAATCATTCTTATGGCCAGGTATATCTCTGTTGA 601  
QY 1820 ACTCAATCTTTTCACTGGAGAAATTAATCTTC 1849  
DB 602 GTGGAGTATGATATGGAGGCAACTCC 631

RESULT 10  
CF474786 700 bp mRNA linear EST 05-SEP-2003  
LOCUS RTW2\_7\_B11\_g1\_A021 Well-watered loblolly pine roots WM2 Pinus taeda cDNA clone RTW2\_7\_B11\_A021 5', mRNA sequence.

DEFINITION  
CF474786  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Pinus taeda (loblolly pine)

REFERENCE  
AUTHORS  
An EST database from well-watered loblolly pine (Pinus taeda) roots  
Unpublished (2003)  
Other ESTs: RTW2\_7\_B11.bl.A021  
Contact: Cordonnier-Pratt MM  
Laboratory for Genomics and Bioinformatics  
The University of Georgia, Department of Plant Biology  
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
Tel: 706 542 1860  
Fax: 706 583 0210  
Email: mmpratt@uga.edu

TITLE  
JOURNAL  
COMMENT  
RNA prepared and library constructed by W. Walter Lorenz, School of

2044 ATCATCGACCAACGCACTGAAGGAATTGAATTGGAGTT 2081  
 |||||  
 659 ATGTCGATTAACCTTAACCAAGGAAGTGAATTGGAGTT 696  
 |||||

RESULT 11  
 CF477103  
 LOCUS  
 DEFINITION  
 RTW3\_5\_A06.g1\_A022 Well-watered loblolly pine roots WM3 Pinus  
 taeda cDNA clone RTW3\_5\_A06\_A022 5', mRNA sequence.  
 CF477103  
 VERSION  
 CF477103.1 GI:34505972  
 EST.  
 SOURCE  
 Pinus taeda (loblolly pine)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus;  
 1 (bases 1 to 745)  
 REFERENCE  
 Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
 Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
 Neale, D.  
 TITLE  
 An EST database from well-watered loblolly pine (Pinus taeda) roots  
 JOURNAL  
 Unpublished (2003)  
 COMMENT  
 Other ESTs: RTW3\_5\_A06.bl\_A022  
 Contact: Cordonnier-Pratt WM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of  
 Forestry, University of Georgia; plant material prepared at the  
 University of Florida; sequencing done in the Laboratory for  
 Genomics and Bioinformatics, University of Georgia. Sequence ends  
 have been trimmed to exclude vector and regions below Phred quality  
 16. Three-prime sequences are presented as their reverse complement  
 and have been trimmed to exclude polyA.  
 Seq primer: JENREV (CAGGAACAGCTGACCT).  
 Location/Qualifiers  
 1..745  
 /organism="Pinus taeda"  
 /mol\_type="mRNA"  
 /accession="CCONES"  
 /db\_xref="taxon:3352"  
 /clone="RTW3\_5\_A06\_A022"  
 /lab\_host="DH10B-T1 phage-resistant E. coli"  
 /clone\_lib="Well-watered loblolly pine roots WM3"  
 /note="Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The  
 library was prepared from polyA+ RNA from loblolly pine  
 (Pinus taeda) roots watered to pot capacity every other  
 day. Pre-dawn water potential remained -0.3 MPa +/-0.1.  
 Roots were harvested for RNA isolation. Double-stranded  
 cDNA was cloned unidirectionally into pSL180. Inserts  
 excised with EcoRI (5' end) and XhoI (3' end)."

FEATURES  
 source  
 Query Match 8.5%; Score 190.4; DB 14; Length 745;  
 Best Local Similarity 54.8%; Pred. No. 4.5e-44;  
 Matches 399; Conservative 0; Mismatches 326; Indels 3; Gaps 1;  
 1449 AATTGCTATGCCAAGACTTCTTGGCTCGGAGTTATTTAGACGATCTTTACGACACCCA 1508  
 |||||  
 1 ACTCGGCTTTGCCAAATCTCTGTCATATCATCACCGTTCTTGATGATATGTACGACACCTT 50  
 |||||

ORIGIN  
 Query Match 8.5%; Score 190.4; DB 14; Length 745;  
 Best Local Similarity 54.8%; Pred. No. 4.5e-44;  
 Matches 399; Conservative 0; Mismatches 326; Indels 3; Gaps 1;  
 1449 AATTGCTATGCCAAGACTTCTTGGCTCGGAGTTATTTAGACGATCTTTACGACACCCA 1508  
 |||||  
 1 ACTCGGCTTTGCCAAATCTCTGTCATATCATCACCGTTCTTGATGATATGTACGACACCTT 50  
 |||||

1509 CGGATCTCGGAGTCTTAATTTGTTCTCTGAAGCGGTCGGAAGTGGATCTCTGT 1568  
 |||||  
 61 CGGAACACTCGAGAGCTCGAATCTTTACAGCTGCAATTAAGAGATGGATCCGTCGC 120  
 |||||

1569 GCTGGATAGCTTGGGATTAATCAGTTGAAAGTTTGTCTTCTAGGCGTGTACACACAGT 1628  
 |||||  
 121 GACAGAGTGCCTTCCAGATATATGAAGGAGTTTACATGATAG---TTTACACACTGT 177  
 |||||

1629 GAATGATTTGGAAAGATCGACTCAAGGAACCAAGCCGCTGATGTCTGGGCTATCTTCG 1688  
 |||||  
 178 AAATGAAATGTCTCAGGAGGACACCAAGGCTCAAGCCGAGACACCGCTCAACTATGTTCG 237  
 |||||

1689 AAAAGTATGGGAGGCTTGTCTCGCATCGTATACCAAGAGCCGAATGTGTCGGCAGCAA 1748  
 |||||  
 238 ACAGGCTTGGGAGGATATATTTGATTCGTATATGCAAGAGCAAGTGGATCGCAGTGG 297  
 |||||

1749 GTATGTCGAGCAATTCAGGAATATGTGGAATAATGCAAAAGTGTCCATAGCACTTGGAC 1808  
 |||||  
 298 TGAGGTGCCAAACATTTGAGGAGTACTACGAGAACCGGAAATTTAGCTCTGGTCATCGCT 357  
 |||||

1809 AGTCGTACTAACTCAATCTTTTCACTCGAGAAATTTACTTCTCGATTTACATTTTACAGCA 1868  
 |||||  
 358 ATCGGCAATTCGAACCCATTTCTGACACCGCATCCCTTTCTTGAGCAGCTCTCAAGGA 417  
 |||||

1869 AGTAGACCTTCGGTCCAAATTTCTGCATCTTGTGTCTTTGACTGGAGCACTTAATCAATGA 1928  
 |||||

418 AGCTGACATTCATCGAGCTCAATGACTTGGCATCTGCCATTTCTTGCATTCAGGAGGA 477  
 |||||

1929 CACCAAGACTTACAGCGCCGAGAGAACCGTGTGTAATTTGTTTCCAGCGTACAGTGTCTA 1988  
 |||||

478 TAGCGCTGTCTACAGCGCGACAGGCGCCGTGGAGAGAAAGCTTCGTGTATATCTTCTTA 537  
 |||||

1989 CATGAGGGAATTCGGAGTGCACAGAGNAGAGCTCTTAAGTCATGTTTTATGTATCAT 2048  
 |||||

538 TATGAAGACAAATCTCTGGAGCAACGGAAGAAGATGCTCTCAATCATATCAACGCCATGAT 597  
 |||||

2049 CGACACGCACTGAAGGAATTTGAGTGTGGCAACCCAGCGAGCAATGCCCAT 2108  
 |||||

598 CAGTGATGTATTAAAGGATTAATTTGGAGCTTCTCAACCAACAGCAGCGTTCCTCAT 657  
 |||||

2109 GTGTGTGAGAAGACTCTCTTCAACACTCAAGAGTATGACAGCTGTTTTATATGTATCAG 2168  
 |||||

658 ATCTGCCAAAACATGCTTTTGACATTTAGCAGAGCTTTCCATTATGGCTACAAAATATCG 717  
 |||||

2169 AGATGGCT 2176  
 |||||

718 AGATGGCT 725  
 |||||

CF401916 696 bp mRNA linear EST 29-AUG-2003  
 RTW1\_15\_B05.g1\_A015 Well-watered loblolly pine roots WM1 Pinus  
 taeda cDNA clone RTW1\_15\_B05\_A015 5', mRNA sequence.  
 CF401916  
 ACCESSION  
 CF401916.1 GI:34360333  
 EST.  
 SOURCE  
 Pinus taeda (loblolly pine)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus;  
 1 (bases 1 to 696)  
 REFERENCE  
 Pratt, L., Cordonnier-Pratt, M.-M., Lorenz, W.W., Dean, J.,  
 Gebremedhin, M., Dervinis, C., Martin, T., White, T., Davis, J. and  
 Neale, D.  
 TITLE  
 An EST database from well-watered loblolly pine (Pinus taeda) roots  
 JOURNAL  
 Unpublished (2003)  
 COMMENT  
 Contact: Cordonnier-Pratt WM  
 Laboratory for Genomics and Bioinformatics  
 The University of Georgia, Department of Plant Biology  
 Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA  
 Tel: 706 542 1860  
 Fax: 706 583 0210  
 Email: mmpratt@uga.edu  
 RNA prepared and library constructed by W. Walter Lorenz, School of  
 Forestry, University of Georgia; plant material prepared at the  
 University of Florida; sequencing done in the Laboratory for  
 Genomics and Bioinformatics, University of Georgia. Sequence ends  
 have been trimmed to exclude vector and regions below Phred quality  
 16. Three-prime sequences are presented as their reverse complement  
 and have been trimmed to exclude polyA.

us-10-041-007-38.rst

Thu Aug 26 08:27:54 2004

```

FEATURES
  source
    Seq primer: JENREV (CAGGAACAGCTATGACC) .
    Location/Qualifiers
      1..696
        /organism="Pinus taeda"
        /mol_type="mRNA"
        /strain="CCLONES"
        /db_xref="taxon:3352"
        /clone="RTW1 15 B05 A015"
        /lab_host="DH10B-T1 Phage-resistant E. coli"
        /clone_lib="Well-watered loblolly pine roots W1"
        /notes="Vector: pSL180; Site 1: EcoRI; Site 2: XhoI; The library was prepared from polyA+ RNA from loblolly pine (Pinus taeda) roots watered to pot capacity every other day. Pre-dawn water potential remained -0.3 MPa +/- 0.1. Roots were harvested for RNA isolation. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts excised with EcoRI (5' end) and XhoI (3' end)."
ORIGIN
  Query Match      8.4%; Score 188.8; DB 14; Length 696;
  Best Local Similarity 55.7%; Pred. No. 1.3e-43;
  Matches 38; Conservative 0; Mismatches 302; Indels 3; Gaps 1;
  QY 1425 TGAGCCAGAAATTCCTGCTGTAGAAATTCCTATGCCAAGACTTCTTGCCTCGCAGTTAT 1484
  DB 12 TGATCTAAACATTTTCGGTTTCGACTAGGATTTGGTAAATAAGTCATATGATCAGAT 71
  QY 1485 TCTAGACGATCTTTACGACACCCACGATCTCTGGATGATCTTAAATTTGTTCTCTGAAGC 1544
  DB 72 TCTCGACGATCTACGACACTTCGGAACAATGGAGAGCTCGAATCTTTAACCCGAGC 131
  QY 1545 GGTCGGAAGATGGGATATCTCTGTCTGGATAGGTTTCGGATATCATGTTGAAAGTTG 1604
  DB 132 GTTTAAGAGATGGGATCGCTCTTCGATAGAGTGTCTTCGATATAT---ATGAAGAGGT 188
  QY 1605 CTTCTAGGCTGTACACACAGATGAATGGATTTGGAAAGATGAGCTCAAGGAACAGG 1664
  DB 189 GTACATGCGGTTTACGACAAACATCAGCAATGGCAGAGGCGCAGAAATTCAGG 248
  QY 1665 CCGTATGCTGGGCTATCTTCGAAAGATATGGAGGGGTTGCTCGCATGTATACAA 1724
  DB 249 CTGGATACAGTCAGCTATGCTCGAAATCTTGGGAGGCTTTTATTTGCTGTATATACA 308
  QY 1725 AGAGCCGAAATGGTCGGCAGCAAGATATGTCGCGACATTCACGCAATATGTGAAATGC 1784
  DB 309 AGAGCCGAAATGGATTTCCAGTGGTTATCTTCCACGTTTCGACGAGTACCTCGAGATGG 368
  QY 1785 CAAAGTGTCCATAGCACTTGCACAGATCGTACTAAATCTCAATCTTTTTCACCTGGAGAAAT 1844
  DB 369 GAAGTTCAGCTTCGGCTCTCGCATACCAACGCTCGAACCACATGCTGACTTTTGGGGTTTCC 428
  QY 1845 ACTTCCTGATTAATTTTACGAGATGACCTTCGGTCCAAATTTCTGCATCTTTGTGTC 1904
  DB 429 TCTTCGCGCTCGAATCTCGAGAAATGACATTTCCATCGAATTCATGATTTGATG 488
  QY 1905 TTGACTGGACGCAATTAATCAATGACACCAAGACTTACAGGGCGGAGAGAAACCGTGGTGA 1964
  DB 489 TGGCATCTCTCGATGAAAGTGAACACTCAATGCTACAGGCTGACAGGGCGGCTGGAGA 548
  QY 1965 ATTGGTTTCCAGCGTACAGTGTACATGAGGGAATTCGGAGTGCACAGAGGAAGAC 2024
  DB 549 AGAAGCTTCGGCGGTATCGTGTATATGAAAGACCATCTCTGGAAATACAGAGGAATGC 608
  QY 2025 TCTAAGTCATGTTTATGTTATCATCGACACGCACTGAAGGAATGAAATTTGGAGTTGGC 2084
  DB 609 TGTCAATCAAGTCAATGCTATGTTGATTAATCTTAACAGGAACTGAATTTGGGAGTTACT 668
  QY 2085 CAACCCGAGCAGCAATGCCCATTTGTT 2112
  DB 669 TAGACCCGACAGCGGTTTCCCATCTCT 696
  RESULT 13
  
```

```

BX677599 581 bp mRNA linear EST 28-OCT-2003
LOCUS BX677599 RN Pinus pinaster cDNA clone RN41G08, mRNA sequence.
DEFINITION BX677599
ACCESSION BX677599
VERSION BX677599.1 GI:38011551
KEYWORDS EST.
SOURCE Pinus pinaster
ORGANISM Pinus pinaster
REFERENCE
  1 (bases 1 to 581)
  Frigerio,J. and Plomion,C.
  Identification of water-deficit responsive genes in Maritime pine
  (Pinus pinaster Ait.) using an EST approach
  Unpublished (2002)
  Contact: Frigerio JM
  Genetique et Amelioration 69
  INRA
  route d'Arcachon 33612 Cestas CEDEX France
  Email: Frigerio@pierrot.inra.fr
  Email: Frigerio@pierrot.inra.fr
  Seq primer: T3
  Location/Qualifiers
    1..581
      /organism="Pinus pinaster"
      /mol_type="mRNA"
      /db_xref="taxon:71647"
      /clone="RN41G08"
      /tissue_type="root"
      /dev_stage="6 weeks old seedling"
      /lab_host="SOLR"
      /clone_lib="SOLR"
      /note="Vector: Uni-ZAP XR; ecotype: Landes; The library was made from the roots of 6 weeks old seedlings grown in hydroponic conditions. A mixture of genotypes were used. Oligo-dT primed cDNA was directionally cloned into the EcoRI-XhoI lambda-ZAP vector arms and mass-excised to form a pBluescript phagemid"
ORIGIN
  Query Match      8.3%; Score 186; DB 13; Length 581;
  Best Local Similarity 64.9%; Pred. No. 7.7e-43;
  Matches 292; Conservative 0; Mismatches 155; Indels 3; Gaps 1;
  QY 1331 TTGTGAGTGTGGAGAAATCGGATTCATGATCTTACATTCACCGCCAGCGCGCTG 1390
  DB 135 TTCAAAGTGTGGAGATCATCCGGTTTCGGGAGCTCAATTTCAACGCGACCGTGTGA 194
  QY 1391 TGGAAATGATTTCTCAGTGGCGGTAGTATGTTTGGAGCCAGAAATTCGCTGTGTAGAA 1450
  DB 195 CTGAATATATTTCTCAATCGGCTCTCCCTGTTTGAACGAGAGCTCGCTTGTGTAGAG 254
  QY 1451 TTGCTATGCGCAAGACTTTCCTCGCGAGTTTATCTAGACGATCTTTACGACCCACG 1510
  DB 255 CCATTTATACGAAATCTACCATCTTGATGTGTCATTTTAGACGACCTATATGACGACATG 314
  QY 1511 GATCTCTGATGATCTTAAATTTGTTCTCGAAGCGGTCCGAAGATCGGATATCTGTGTC 1570
  DB 315 CATCTTTGAAGACATCAAGCTGTTACAGAGAGCAGTTAAAGATGGGATCTGTTCTGT 374
  QY 1571 TGGATAGCGTTCCGGGAATATCAGTTGAAAGTTTCTTCTAGGGCTGTACACACAGTGA 1630
  DB 375 TAGACCGAATGCCAG---AACACATCAAAACATGCTACCTGGGTTTGTACATCTTGTGA 431
  QY 1631 ATGGATTTGGAAAGATGACTCAAGGAACAAGCGGTGATGTGCTGGGCTATCTTCGAA 1690
  DB 432 ATGAATAGCAGAAAGAGGGCGGACAGACAGGCTGCTGACGTACTAGGCTACATTAGAA 491
  QY 1691 AAGTATGGAGGGCTTCTCGCATCTGTATACCAAGAGCCGAAATGCTCGGACCAAGT 1750
  DB 492 ATTATGGGAGATCCTGTTTGAAGCATATACGAAGAAGCAGAAATGCTCGGAAGCTGCTT 551
  QY 1751 ATGTGGCAGATTCACAGGAATATGTGAAA 1780
  
```

```
Db      552  ATGTGCTCTCTTCAATGAATACATAGAGA 581
|||||
BU971902      624 bp      mRNA      linear      EST 22-OCT-2002
HB20A061 BC Hordeum vulgare subsp. vulgare cDNA clone HB20A06
5-PRIME, mRNA sequence.
ACCESSION   BU971902
VERSION     BU971902.1 GI:24222695
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare
ORGANISM    Hordeum vulgare subsp. vulgare
REFERENCE   1 (bases 1 to 624)
AUTHORS    Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.
TITLE      Barley ESTs from developing seeds
JOURNAL    Unpublished (2002)
COMMENT    Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 20 row: A column: 6
Seq primer: M3rev.
Location/Qualifiers
1..624
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="barke"
/sub_species="vulgare"
/db_xref="GABI:236738"
/db_xref="taxon:112509"
/clone="HB20A06"
/issue_type="developing caryopsis"
/dev_stage="8-15 DAP (days after pollination)"
/lab_host="Xl10-Gold"
/clone_lib="BC"
/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); developing caryopsis, 8-15 DAP(days after pollination) Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

ORIGIN
Query Match      8.2%; Score 184.4; DB 13; Length 624;
Best Local Similarity 59.5%; Pred. NO. 2.4e-42;
Matches 368; Conservative 0; Mismatches 241; Indels 9; Gaps 3;

QY 679  GATCTATTGAAGCCTCTGGGTCTCGATACAGTTGAACGCTTGGGATCGATCGCTAT 738
|||||
Db 11  GATCTCTTTGAGCGCATCTGGTCCGTCGATCGGTTAGAGCGTCTTGGAAATCTCGCGCTAT 70
|||||
QY 739  TTCAAGAAAGAAATCAAGAAATCTCTGGATTACGTTTATAGGTACTGGGACGCCGAAAGA 798
|||||
Db 71  TTCAGCAAGAAATCAAGCAGCTCTGGATTATGTTTCAGAGGACATGG---ACTGATGAG 127
|||||
QY 799  GCGGTGGGATGGCAAGATGCATCTCTTCTGATGTGATGACACTGCCATGGGCTCTT 358
|||||
Db 128  GGGATTTGCTGGCGAGGAACCTCCACTGTAATAGACGTGATGATACATCCATCCGATTC 187
|||||
QY 859  AGAATCCTGAGACTTCATGATACAAATGATCTCTTCTGATGTTCTGAGAAATTCAGAGAC 318
|||||

Db      188  CGCTCTCGGGCTTCATGATACGATGCTCTCCCAAGTGTATTTCGGAAGTT---TGAG 244
QY 919  GAGAAAGGAGACTTCTTTTGGCTTGGCGGTCAACGCAAAATTGGTGTGACCGATAATCTT 978
|||||
Db 245  AAGGACGGGGAGTCTTCTTGTTCGTGGGCAATCAACAACGAGTCACTGGGATGATC 304
|||||
QY 979  AACCTTTATAGATGTTTCAAGATATGTTTTCGGGAGAAAGATAATGGAGAAGCTTAAG 1038
|||||
Db 305  AACCTGAACAGGGCCCTCTCAGGTAAGGTTCCCTGGAGAGGACTTGTTCACGCGTGCAAGG 364
|||||
QY 1039  ACCTTCACATACAAATCATCTCCAAATGCTCTTGGCAAAAACAACGCAATTTTGATAAGTGG 1098
|||||
Db 365  AGATTCCTCGTATGAGTTCCTTAGAGAAAGGAGCCCATCGGACGATTCGAGACAAATGG 424
|||||
QY 1099  GCTGTCAAGAAGGATCTCTCGAGAGGTGGAGTATGCTATATAAGATATCGTGGCAGAGA 1158
|||||
Db 425  ATCATTGCTAAGGATCTACACGCGAGGTAAATAATACACTGACCTCCCATGGTATGCA 484
|||||
QY 1159  AGTATGCCAAGATTGGAGGCAAGAGTTACATAGACCAATTT---GGATCAATATGATGTC 1215
|||||
Db 485  AGCTTACCGCGTGTAGAAGCAAGAGTCTACCTAGATCAATATATGGCGTGTATATGATGTC 544
|||||
QY 1216  TGGCTGGGGAAGACTGCTGTATAGATGCTATATGTCAGCAACGAAATAATTTTGGAGCTG 1275
|||||
Db 545  TGAATTTGGGAAGACACTCTACAGGATGCTACTTTGTGAACAACAACACCTATCTTGAGTTG 604
|||||
QY 1276  GCCAAATTGGACTTCAAT 1293
|||||
Db 605  GCAGAGCGTGATTTCAAT 622
|||||

RESULT 15
LOCUS      CA223665
DEFINITION CA223665 688 bp mRNA linear EST 25-SEP-2003
SCJFFLIC04F02.g FL1 Saccharum officinarum cDNA clone SCJFFLIC04F02
5', mRNA sequence.
ACCESSION   CA223665
VERSION     CA223665.1 GI:35280237
KEYWORDS    EST.
SOURCE      Saccharum officinarum
ORGANISM    Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Saccharum.
REFERENCE   1 (bases 1 to 688)
AUTHORS    Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
TITLE      The libraries that made SUCEST
JOURNAL    Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT    Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: C04 row: F column: 02
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1..688
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCJFFLIC04F02"
/clone_lib="FL1"
/lab_host="DH10B"
/note="Organ: Inflorescence at beginning of development
(lcm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;
An unidirectional cDNA library generated from
[Inflorescence at beginning of development (lcm-long)].
cDNA was prepared from polyA+ mRNA using SuperScript
```

Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <http://sucest.lad.ic.unicamp.br/public>

ORIGIN		Query Match	8.2%;	Score 184.2;	DB 14;	Length 688;
		Best Local Similarity	57.1%;	Pred. No. 2.9e-42;		
		Matches 396;	Conservative 0;	Mismatches 288;	Indels 9;	Gaps 3;
Qy	751	ATCAAGAATCTCTGGATTACGTTTATAGTACTGGACCCGGAAGAGCGCTGGGATGG	810			
Db	2	ATTAAGCAATGTCATGGACTATGTGAACAGGCACCTGG---ACTGAAGATGGGATTTGCTGG	58			
Qy	811	GCAGATGCAATCCTATTCTCTGATGTCGATGACACTGCCATGGGTCTTAGAATCCTGAGA	870			
Db	59	GCTAGGGACTCTACTGTAAAGATGTGNGATGACACAGCTATGCTTCCGACTACTAAGG	118			
Qy	871	CTTCATGGATACATATGATCTTTCAGATGTTCTGGAGAAATTCAGAGACGAGAAAGGAGAC	930			
Db	119	CTACATGATACATATGCTCCCAAGTGTGTTTAAGAAATTT---TGAGAAAGATGGAGAG	175			
Qy	931	TTCTTTTGTGCTTGGCGGTCAAAAGCAATTTGGTGTGACCGGATATCTTAACCTTTTATAGA	990			
Db	176	TTCTTTTGTGTTTGTGGGGCAATCAACTCAAGCCGTCACTGGGATGTACAACCTCAACAGA	235			
Qy	991	TGTTTCAAGTATGCTTTTCCGGGAGAAAGATAATGGAAGAGCTAAAGACCTTCACCTACA	1050			
Db	236	GCCTCTCAATATAGCTTTTCAAGGGGAGGATGATTAACACCGTGTAGGACTTTCTCATAT	295			
Qy	1051	AATCATCTCCAAATATGCTCTTTGCCAAACAAACGCAATTTGATAGTGGGCTGTCAAGAG	1110			
Db	296	GAGTTTCTCAGACAAAGAGAGCCCAAGGCATGCTCCGTGATAAATGGATCATCGCGAAG	355			
Qy	1111	GATCTTCTGAGAGGTGGAGTATGCTATAAGTATCCGTGGCATAGAAATGATGCCAAGA	1170			
Db	356	GATCTAGCTGGCGAGGTACAATATACACTAGACTTCCCTTGGTATGCAAGCTTGCCTCGT	415			
Qy	1171	TTGGAGCAAGAAGTTACATAGAGCAATTTGGATCAA---ATGATGTCTGGCTGGGGAAG	1227			
Db	416	GTGAGGCAAGAACCTATCTAGATCAGTATGGCGTAAAGATGATGTTGGATTGGAAG	475			
Qy	1228	ACTGTGTATAGATGCTATATGTGAGCAACGAATAATTTGGAGCTGGCCAAATTTGGAC	1287			
Db	476	ACCTATACAGGATGCCCTTTGTGATAACGACATATATCTTGATTTGGCAATAATGGAT	535			
Qy	1288	TTCAATATGGTGCAGGCTTTACCCAAAGAGAGACTCAACACATTTGTGAGCTGGTGGAGA	1347			
Db	536	TTCAACCGTTGCCAAGCTCTACATCAGCTTGATGTGATGAGCTTCAATGTGGTACATA	595			
Qy	1348	GAAATCGGATTCAAATGATCTTACATTTCAACCCGAGCGGCTGTGGAAATGATTTCTCA	1407			
Db	596	GAGAAATGCTTTGACACTTTTGGAGTGCAACCGCAAGATGTTTAAAGAGCTTATTTTAA	655			
Qy	1408	GTGGCGGTTAGTATGTTTGAGCCAGCAATTCGCT	1440			
Db	656	GCTGGCTCTTGCAATTTATGAACCTTAGTCTGCT	688			

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: August 17, 2004, 20:53:08 ; Search time 46.5386 Seconds  
(without alignments)  
4529.153 Million cell updates/sec

Title: US-10-041-007-39  
Perfect score: 3972  
Sequence: 1 MFQSMGDTNPAYDTAWV.....GISDKEMKHVSRITLFDPA 746

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3972	100.0	873	ABU09779	Maidenhair
2	2735.5	68.9	868	AAy08570	Delta-sel
3	2735.5	68.9	868	AAy08563	Grand fir
4	2713.5	68.3	868	AAy08570	Grand fir
5	2050.5	51.6	782	AAW85703	Grand fir
6	2050.5	51.6	782	AAy08570	Grand fir
7	2050.5	51.6	782	AAW85703	Grand fir
8	2050.5	51.6	782	AAy08570	Grand fir
9	2050.5	51.6	817	AAy08567	E-alpha-b
10	2050.5	51.6	817	AAy08566	Grand fir
11	2050.5	51.6	817	AAy08562	Grand fir
12	2050.5	51.6	817	AAy08561	Grand fir
13	2050.5	51.6	817	ABU09781	White fir
14	2049.5	51.6	817	AAy08568	E-alpha-b
15	1901	47.9	862	AAW31655	Pacific y
16	1901	47.9	862	AAy08562	Grand fir
17	1901	47.9	862	AAy08561	Grand fir
18	1901	47.9	862	AAy08562	Grand fir
19	1363.5	34.3	815	AAy08568	E-alpha-b
20	1363.5	34.3	815	AAW31655	Pacific y
21	1320	33.2	802	AAy08562	Grand fir
22	1320	33.2	802	AAy08561	Grand fir
23	1292.5	32.5	823	AAy08562	Grand fir
24	1255	31.6	581	AAy08569	Delta-sel
25	1252	31.5	580	AAy06571	Delta-sel

26	1252	31.5	581	2	AAy06570	Delta-sel
27	1252	31.5	581	2	AAy06563	Grand fir
28	1252	31.5	581	3	AAy08564	Grand fir
29	1215.5	30.6	630	4	AAW85704	Grand fir
30	1206	30.4	577	2	AAW85704	Grand fir
31	1206	30.4	577	4	AAW85704	Grand fir
32	1194.5	30.1	637	2	AAW85702	Limonene
33	1194.5	30.1	637	3	AAy08569	Grand fir
34	1194.5	30.1	637	4	AAW85702	Limonene
35	1193	30.0	637	4	AAW85702	Limonene
36	1173.5	29.5	627	2	AAW85710	Grand fir
37	1173.5	29.5	627	4	AAW85710	Grand fir
38	1173.5	29.5	627	5	ABW79395	Taxadiene
39	1167.5	29.4	627	2	AAW85700	Myrcene s
40	1167.5	29.4	627	3	AAy08564	Grand fir
41	1167.5	29.4	627	4	AAW85700	Myrcene s
42	1151.5	29.0	593	2	AAy06573	Gamma-hum
43	1151.5	29.0	593	2	AAy06572	Gamma-hum
44	1151.5	29.0	593	2	AAy06564	Grand fir
45	1151.5	29.0	593	3	AAy08565	Grand fir

## ALIGNMENTS

RESULT 1  
ABU09779

ID ABU09779 standard; protein; 873 AA.

XX ABU09779;

XX 06-AUG-2003 (first entry)

XX Maidenhair tree levoipimaradiene synthase.

XX Levopimaradiene synthase; ginkgolide; enzyme; Escherichia coli;

XX Candida albicans; Kluyveromyces fragilis; Saccharomyces; maidenhair tree.

XX Ginkgo biloba.

XX US2002164736-A1.

XX 07-NOV-2002.

XX 07-JAN-2002; 2002US-00041007.

XX 05-JAN-2001; 2001US-0259881P.

XX (MATS/) MATSUDA S P T.

XX (SCHE/) SCHEPMANN H G.

XX Matsuda SPT, Schepmann HG;

XX WPI; 2003-352584/33.

XX New purified and isolated nucleic acid sequence encoding a levoipimaradiene synthase useful for producing pure ginkgolides in cells.

XX Claim 20; Fig 4; 37pp; English.

XX The invention relates to a purified and isolated nucleic acid sequence encoding a levoipimaradiene synthase. The nucleic acid is used for the preparation of a ginkgolide in a cell, an expression vector, a unicellular organism (e.g. Saccharomyces, Escherichia coli, Candida albicans or Kluyveromyces fragilis), a host cell, a plant (particularly Ginkgo biloba), a host cell, a transgenic plant or a seed of the transgenic plant, for the biosynthesis of ginkgolides and in molecular biology, molecular genetics, and organic chemistry. The nucleic acid is capable of producing large amounts of pure ginkgolides in a cost-effective manner, particularly in an organism capable of a high-yield ginkgolide-producing system. This sequence represents a levoipimaradiene synthase polypeptide of the invention

XX

signal transduction pathway; bile acid; affinity purification; photoreceptor; enzymatic synthesis; nutrient supplement; immunological reagent.

[illegible]

RESULT 2	
AAAY90858	
ID	AAAY90858 standard; protein; 869 AA.
XX	
XX	
AC	AAAY90858;
XX	
XX	
DT	25-AUG-2000 (first entry)
XX	
DE	Grand fir abietadiene synthase protein sequence SEQ ID NO:56.
XX	
KW	Synthase; protein co-ordinate data; active site; modification; terpenoid
KW	3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW	3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;
KW	isoprenoid; breeding programme; fragrance; flavour; pheromone;
KW	defensive agent; pigment; antitumour; steroid hormone;
KW	defensive agent; pigment; antitumour; steroid hormone;

KW	signal transduction pathway; bile acid; affinity purification;
KW	photoreceptor; enzymatic synthesis; nutrient supplement;
KW	immunological reagent.
XX	
XX	Abies grandis.
XX	WO200017327-A2.
XX	
XX	30-MAR-2000.
PD	
XX	17-SEP-1999; 99WO-US021419.
XX	
XX	18-SEP-1998; 98US-0100993P.
PR	22-APR-1999; 99US-0130628P.
PR	23-AUG-1999; 99US-0150262P.
XX	
PA	(KENT ) UNIV KENTUCKY RES DEPT.
PA	(SALK ) SALK INST BIOLOGICAL STUDIES.
XX	
PI	Chappell J, Manna KR, Noel JP, Starke CM;
XX	
XX	WPI; 2000-292839/25.
DR	N-PSDE; AAA39937.
DR	
XX	
XX	Novel terpene synthase enzymes, useful for producing terpene
PT	hydrocarbons, e.g. fragrances or antitumor agents, are derived from known
PT	enzymes by specific amino acid alterations.
XX	
XX	Claim 112. Page 443-445; 450pp; English.

The present invention describes an isolated terpene synthase (I) comprising a region with at least 20% identity to region 265-535 of a 548 amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha-carbon atoms (alphaC) that have interatomic distances, between each other, within tabulated ranges, have a centre point (within a sphere of radius 2.3 Angstrom) within tabulated ranges, and have an ordered arrangement of R groups (defining aa side chains), excluding specific tabulated arrangements (tables given in the specification). (I), and related enzymes, are used to produce a wide range of terpenoids (e.g. cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances, flavours, pheromones, defensive agents, pigments, antitumour agents, components of signal transduction pathways, precursors of steroid hormones and bile acids, as photoreceptors and as co-factor side chains. Some synthases with little or no catalytic activity and as co-factor side chains. encoding them) are used as controls in the analysis of products formed by enzymatic synthesis; as nutrient supplement; for affinity purification of isoprenoids; or to develop immunological reagents or nucleic acids for monitoring expression of terpene synthase or inheritance of the gene in plant breeding programs. The new synthases may produce novel terpene products. AAA38910 to AAA38938 and AAY90831 to AAY90859 represent the classification of the present invention

XX Sequence 868 AA; SO

Query Match	68.9 %;	Score 2735.5;	DB 3;	Length 868;
Best local similarity	67.6 %;			
Predicted similarity	67.6 %;			
Pred. No.	3.8e-237;			

Qy	1	MFQSMGDETPNSAYDTAWARIPSIDSGAPQFPQTLWIILNNQLPDGSGWGEICIFLAY
Qz	126	MFPCWYGETNPNSAYDTAWARIPAVDSGDNHPHPEKVEWILQNLQDXGSGWGGFYFLAY
Qx	185	MFPCWYGETNPNSAYDTAWARIPEKVEWILQNLQDXGSGWGGFYFLAY

[illegible]

DB  
QY

186 DRILATACIIITULMKGEIQVQRGHEFFRAAGAAKALAEI  
121 KSLGLDLPYHLPEISQHOKRKKLQKPTPLNVLNHQTALLYSLEGLQDVVDQOEITNLQ

DB	QY	246	181	24
		KILGLDLFYDLPELKQIIIEKREAKLRPTDVIYALPTITLLYSUEGLQSLVDWQKATRDQ	SRGSEFLSPASTACVFVHHTQNKRCLEHFNFLSKFGDVPVCHYPPLDLPERLWADVTVER	

Db







XX PS  
XX PS  
XX PS  
CC CC Nucleotide sequences encoding myrcene synthase, limonene synthase and  
CC pinene synthase from Grand fir may be incorporated into any organism  
CC (e.g. intact plant, animal, microbe), or derived cell culture that  
CC produces geranyl diphosphate for the production of the aforementioned  
CC enzymes or their products. The sequences when expressed in transfected  
CC cells may also be used for the production or modification of flavour and  
CC aroma properties, improvement of defense capability, and the alteration  
CC of other ecological interactions mediated by myrcene, limonene, pinene,  
CC or their derivatives. In particular they can be used for the production  
CC of plant seeds for the extraction of oil or meal. cDNAs encoding myrcene  
CC synthase, pinene synthase and limonene synthase from Grand Fir were  
CC isolated and sequenced by the following procedure. Based on comparison of  
CC sequences of different plant synthase enzymes, four conserved sequences  
CC were identified for construction of four degenerate primers (AA08646-  
CC 49). One of these primers (AA08648) was an antisense primer and used for  
CC PCR with the other three sense primers. Only a combination of primers  
CC AAX08648 and AAX08649 generated a PCR product which was 110 base pairs in  
CC length. The amplified product was used in a plasmid construct to  
CC transform E. coli Xli-blue cells and the inserts were sequenced. Four  
CC probes (AAX08650-53) were devised from the sequenced insert and were used  
CC to screen a Grand Fir cDNA library. One of these clones is the myrcene  
CC synthase sequence given in AAX08643. The other clones identified are  
CC described in AAX08654-56  
XX  
SQ Sequence 782 AA;  
Query Match  
Best Local Similarity 51.6%; Score 2050.5; DB 2; Length 782;  
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;  
6 GDGET--NPSAYTAWARIPSDGSGAPQPTQLQWLNQPLDGSGBECIFLAYDRV 53  
38 GDGESMITPSAYTAWARVPAIDGSGARPPQPTQVDWILKNQKDGSGWGIQSHFLSDRL 37  
64 LNTLACLLTLKIWNKGDIOVKQGVFVRKMEKMEKDEAD-NHRPSGFVFPVPPAMLEAKS 122  
98 LATLSCVLLKWNVGDQVEQGEIRFKSNLELVKDETDQDSLVTDFTFIFFSLLREAOQS 157  
123 LGLDLVPHLPFTSIQHKQKQKLOKPLNVLHNTQALLYSLEGLQVDVDMQBITNLQSR 192  
158 LRGLGFDYDLPYIHLQTKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 217  
183 DGSFLSSPASTACVFMHTQKRCCLHFLNVLKSGDYVPCYHPLDLFLERLWADTVTERLG 242  
218 DGSFLSSPASTACVFMHTGDAKCLFLNSWMIKFGNFVPCLYPVDLLERLIVDNLVRLG 277  
243 IDRYFKKEIKESLDYVRYWDARGVGWACNPIDVDDTAMGLRIILRHGYNVSSDVLE 302  
278 IYRHFKEIKESLDYVRYWDARGVGWACNPIDVDDTAMGLRIILRHGYNVSSDVLE 336  
303 NPREDKGDFFCPAGQTOIGVTDNLNLYRCQVCPGCKIMEBAKTFTTNHQLNALAKNA 362  
337 NFDANGKFCSTGQFNKQVAMNLNLYRASQAPGGENILDEAKSPATKYLREALKSET 396  
363 PDKMAVKDLPGEVYAIKYPHRSMPLEARSYIEQFGSNDVLMGKTYKMLYVNEKY 422  
397 SSANWNNKLSQIKYALKTSWASVPRVEAKYCVYRDPYARIKACVYKLPVYVNEKF 456  
423 LEAKLDFFNNQALHOKETQHVSWWRESGNDLTTTRQRPVEMYSVAVSMPEPFAAC 482  
457 LEIGKLDFFNIQSIHQEBKMNVTWSPRDSGLPLFTFARERPLEFVLAAGTYEPYAKC 516  
483 RIAYAKTSCAVLTDLDLYTHGSLDGLKLPSEAVRWDISVDSVDNQLKVCFLGYNT 542  
517 RFLFTKVALQTVLDDMDYDTGLDELKLPTEAVRRWDLSTENLFD-YMKLCYQIYDI 575  
543 VNGFGKDGKLEQGRDVLGYLRKWEGLLASYTKAEWSAAKYVPTFNEVYENAKVIALA 502  
576 VHEVAWEAKESQRELVSFRKQWEDYLLGYEEAEWLAAYVPTLDEIKNGITSIGQR 635

Example 1; Page 90-92; 121pp; English.

Nucleotide sequences encoding myrcene synthase, limonene synthase and pinene synthase from Grand fir may be incorporated into any organism (e.g. intact plant, animal, microbe), or derived cell culture that produces geranyl diphosphate for the production of the aforementioned enzymes or their products. The sequences when expressed in transfected cells may also be used for the production or modification of flavour and aroma properties, improvement of defense capability, and the alteration of other ecological interactions mediated by myrcene, limonene, pinene, or their derivatives. In particular they can be used for the production of plant seeds for the extraction of oil or meal. cDNAs encoding myrcene synthase, pinene synthase and limonene synthase from Grand Fir were isolated and sequenced by the following procedure. Based on comparison of sequences of different plant synthase enzymes, four conserved sequences were identified for construction of four degenerate primers (AAX08646-49). One of these primers (AAX08648) was an antisense primer and used for PCR with the other three sense primers. Only a combination of primers AAX08648 and AAX08649 generated a PCR product which was 110 base pairs in length. The amplified product was used in a plasmid construct to transform E. coli Xli-blue cells and the inserts were sequenced. Four probes (AAX08650-53) were devised from the sequenced insert and were used to screen a Grand Fir cDNA library. One of these clones is the myrcene synthase sequence given in AAX08643. The other clones identified are described in AAX08654-56

Query Match  
Best Local Similarity 51.6%; Score 2050.5; DB 2; Length 782;  
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;

6 GDGET--NPSAYTAWARIPSDGSGAPQPTQLQWLNQPLDGSGBECIFLAYDRV 53  
38 GDGESMITPSAYTAWARVPAIDGSGARPPQPTQVDWILKNQKDGSGWGIQSHFLSDRL 37  
64 LNTLACLLTLKIWNKGDIOVKQGVFVRKMEKMEKDEAD-NHRPSGFVFPVPPAMLEAKS 122  
98 LATLSCVLLKWNVGDQVEQGEIRFKSNLELVKDETDQDSLVTDFTFIFFSLLREAOQS 157  
123 LGLDLVPHLPFTSIQHKQKQKLOKPLNVLHNTQALLYSLEGLQVDVDMQBITNLQSR 192  
158 LRGLGFDYDLPYIHLQTKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 217  
183 DGSFLSSPASTACVFMHTQKRCCLHFLNVLKSGDYVPCYHPLDLFLERLWADTVTERLG 242  
218 DGSFLSSPASTACVFMHTGDAKCLFLNSWMIKFGNFVPCLYPVDLLERLIVDNLVRLG 277  
243 IDRYFKKEIKESLDYVRYWDARGVGWACNPIDVDDTAMGLRIILRHGYNVSSDVLE 302  
278 IYRHFKEIKESLDYVRYWDARGVGWACNPIDVDDTAMGLRIILRHGYNVSSDVLE 336  
303 NPREDKGDFFCPAGQTOIGVTDNLNLYRCQVCPGCKIMEBAKTFTTNHQLNALAKNA 362  
337 NFDANGKFCSTGQFNKQVAMNLNLYRASQAPGGENILDEAKSPATKYLREALKSET 396  
363 PDKMAVKDLPGEVYAIKYPHRSMPLEARSYIEQFGSNDVLMGKTYKMLYVNEKY 422  
397 SSANWNNKLSQIKYALKTSWASVPRVEAKYCVYRDPYARIKACVYKLPVYVNEKF 456  
423 LEAKLDFFNNQALHOKETQHVSWWRESGNDLTTTRQRPVEMYSVAVSMPEPFAAC 482  
457 LEIGKLDFFNIQSIHQEBKMNVTWSPRDSGLPLFTFARERPLEFVLAAGTYEPYAKC 516  
483 RIAYAKTSCAVLTDLDLYTHGSLDGLKLPSEAVRWDISVDSVDNQLKVCFLGYNT 542  
517 RFLFTKVALQTVLDDMDYDTGLDELKLPTEAVRRWDLSTENLFD-YMKLCYQIYDI 575  
543 VNGFGKDGKLEQGRDVLGYLRKWEGLLASYTKAEWSAAKYVPTFNEVYENAKVIALA 502  
576 VHEVAWEAKESQRELVSFRKQWEDYLLGYEEAEWLAAYVPTLDEIKNGITSIGQR 635

603 TVVLNSI-PFTGELLPDYILQOVDL--RSKFLHLVSLTGLINDTKTYQARRNGELVSS 659  
636 ILLSGLVINDMGQLLSQEALEKVDYPPGRVLTLSLISLADDTKTYKAEKARGELASS 695  
660 VOCVMENPECTEERALSHVYGIIDNALKEINWELANPASNAPLCVRRLLPNTARVMQLF 719  
696 IECYMKDHECTEERALDHIYSILEPAVKELTREFLPK-DDVPFACKMLFEETRVMTVI 754  
720 YMYRDGFGISDKEMKHVSRTLPDPV 745  
755 FKDDGFGVSKLEVKDHIKECLIEPL 780

RESULT 6  
AAY90853  
ID AAY90853 standard; protein; 782 AA.  
XX AAY90853;  
XX AAY90853;  
DT 25-AUG-2000 (first entry)  
XX  
DE Grand fir E-alpha-bisabolene synthase protein sequence SEQ ID NO:46.  
XX  
KW Synthase; protein co-ordinate data; active site; modification; terpenoid;  
KW 3-dimensional coordinate; alpha carbon atom; plant; terpene synthase;  
KW isoprenoid; breeding programme; fragrance; flavour; pheromone;  
KW defensive agent; pigment; antitumor; steroid hormone;  
KW signal transduction pathway; bile acid; affinity purification;  
KW photoreceptor; enzymatic synthesis; nutrient supplement;  
KW immunological reagent.  
XX  
OS Abies grandis.  
XX  
XX WO200017327-A2.  
XX  
PD 30-MAR-2000.  
XX  
PF 17-SEP-1999; 99WO-US021419.  
XX  
XX 18-SEP-1998; 98US-0100993P.  
PR 22-APR-1999; 99US-0130628P.  
PR 23-AUG-1999; 99US-0150262P.  
XX  
XX (KENT) UNIV KENTUCKY RES DEPT.  
PA (SALK) SALK INST BIOLOGICAL STUDIES.  
XX  
XX Chappell J, Manna KR, Noel JP, Starks CM;  
XX  
XX WPI; 2000-292839/25.  
DR N-PSDB; AAA38932.  
XX  
XX Novel terpene synthase enzymes, useful for producing terpene  
PT hydrocarbons, e.g. fragrances or antitumor agents, are derived from known  
PT enzymes by specific amino acid alterations.  
XX  
PS Claim 108; Page 418-420; 450pp; English.  
XX  
CC The present invention describes an isolated terpene synthase (I)  
CC comprising a region with at least 20% identity to region 265-535 of a 548  
CC amino acid (aa) sequence (Ia), given in AAY90831. (I) contains nine alpha  
CC -carbon atoms (alphaC) that have interatomic distances, between each  
CC other, within tabulated ranges, have a centre point (within a sphere of  
CC radius 2.3 Angstrom) within tabulated ranges, and have an ordered  
CC arrangement of R groups (defining aa side chains), excluding specific  
CC tabulated arrangements (tables given in the specification). (I), and  
CC related enzymes, are used to produce a wide range of terpenoids (e.g.  
CC cyclic, acyclic, optionally hydroxylated), useful e.g. as fragrances,  
CC flavours, pheromones, defensive agents, pigments, antitumor agents,  
CC components of signal transduction pathways, precursors of steroid  
CC hormones and bile acids, as photoreceptors and as co-factor side chains.  
CC Some synthases with little or no catalytic activity (and nucleic acids  
CC encoding them) are used as controls in the analysis of products formed by  
CC enzymatic synthesis; as nutrient supplements; for affinity purification

CC of isoprenoids; or to develop immunological reagents or nucleic acids for  
CC monitoring expression of terpene synthase or inheritance of the gene in  
CC plant breeding programs. The new synthases may produce novel terpene  
CC products. AAA38910 to AAA38938 and AAA90831 to AAA90859 represent  
CC sequences used in the exemplification of the present invention  
XX  
SQ

Query Match 51.6%; Score 2050.5; DB 3; Length 782;  
Best Local Similarity 50.8%; Pred. No. 2e-175;  
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;  
QY 6 GDGET--NPSAYDTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGECEIFLAYDRV 63  
DB 38 GDGESMITPSAYDTAWARVPAIDGSAARPQFPQTVDWILKNQLKDGSGWGIQSHFLSDRL 97  
QY 64 LNTLACLLTKIKNKGDIQVQKGVFVRKHEEMKDEAD-NHRPSGFEVVPFAMLDKAS 122  
DB 98 LATLSCLVLLKWNVGLQVEQIEFKSNLELVKDETDQDLSLVTDFEIFFPSLLREAO 157  
QY 123 LGLDLPYHLPFTISQIHQKQKLOKIPNLVNNHQTALLYSLGLEQDVVDWQEIITNLSR 182  
DB 158 LRGLPYDLPYIHLQTKQERLAKLSREIYAVPSPLLYSLEGIQDIVEWERIMEVQSQ 217  
QY 183 DGSFSLSPASTACVFMHTQNKCLHFLNVLKSGDVVPCYPLDLFERLWAVDTVERLG 242  
DB 218 DGSFSLSPASTACVFMHTGDAKCLEFLNSVMKEFGNVPCLYPVLLERLLIVDNI 277  
QY 243 IDRFYKKEIKESLDYVYRYWDAERGVGWACNPIPDVDDTAMGLRIILRHGYNVSSDYLE 302  
DB 278 IYRHFKEIKALDYVYRHNW-ERGIGWGRNLNADLETALGFRLLRLHRYNVSIPAIFD 336  
QY 303 NFRDEKGFDFCFAGQTOIGVTDNLNLYRCSQVCPGKIMEBEAKTFTTNHLQNALAKNA 362  
DB 337 NFKDANGKFCSTGQFNKDVASMLNLYRASQLAPPGENILDEAKSFATKYLREALEKSET 396  
QY 363 FDKWAVKDKLPGEVEYAIKYPWHRSMPLREARSYIEQFGSNDVMLGKTVYKMLYVSNKY 422  
DB 397 SSANNKQNLSEIKYALKTSWHSAPRVEAKRYCQVYRDPYARIKACVYKLPYVNNKFP 456  
QY 423 LELAKLDNMVQALHOKETQHIIVSWWRBESGFNDLFTTRQRPVEMYFVSVAVSMFEPEFAAC 482  
DB 457 LELGKLDNFIIQSIIHQEEMKNVTSFRDSGLPLFTFARERPLEYFLVAAGTYEPQYAKC 516  
QY 483 RIAYAKTSCIAVLDDLYTHGSLDDLKLFSEAVRRNDISVLDSVRDNOLKCYFLGLYNT 542  
DB 517 RELFTKVAQLQTVLDDMYDTYGLDELKLFTEAVRRNDLSFTENLPD-YMKLCYQIYDI 575  
QY 543 VNGFGKDGKLEGRDVLGYLRKWEGLLASYTEKEAESAAYKVPFTFNEYVENAKYSIALA 602  
DB 576 VHEVAWEAEKEQRELVSFRRKWDYLLGYEEAEWLAAYVPTLDYIKNGITSIGOR 635  
QY 603 TVVLNSI-EFTGELLPDYILQQVDL--RSKFLHLSLTGRLINDTKTYQAERNRGLVSS 659  
DB 636 ILLLSGLVMDQGLLSQEALEKVDYPCRRVLTLEINLSRLADDTTKTYKAERKARGLASS 695  
QY 660 VOCYMRNEPECTEERALSIVYGIIDNALKELNWLNPASNAPLCVRRLLNTFARWQOLF 719  
DB 696 IECYMKDHPCECTEERALDHIYSILEPAVKELTREFLKP-DDVPFACKMLFEETRVTWVI 754  
QY 720 YMYRDGFGISDKEMKHVSRTLPDPV 745  
DB 755 FKDGDFGVSKLEVKDHIKECLIEPL 780

RESULT 7  
AAB69373  
ID AAB69373 standard; protein; 782 AA.  
XX  
AC AAB69373;  
XX  
DT 30-APR-2001 (first entry)  
XX

DE Grand fir abietadiene synthase protein fragment SEQ ID NO: 16.  
XX Monoterpene synthase; grand fir; cancer; (-)-camphene synthase;  
KW myrcene synthase; (-)-limonene synthase; (-)-pinene synthase;  
KW terpinolene synthase; insect resistance; nutrition.  
XX  
OS Abies grandis.  
XX  
PN WO200107565-A2.  
XX  
PD 01-FEB-2001.  
XX  
PF 24-JUL-2000; 2000WO-US020264.  
XX  
PR 26-JUL-1999; 99US-00360545.  
XX  
XX (UNIW ) UNIV WASHINGTON STATE RES FOUND.  
XX  
XX Steele CL, Bohlmann J, Croteau RB, Phillips MA;  
XX  
XX WPI; 2001-182782/18.  
XX  
XX N-PSDB; AAF73382.  
XX  
XX New nucleic acid encoding monoterpene synthases, for increasing terpene  
XX synthesis in plants, e.g. for increasing resistance to pests or for  
XX treatment of cancer.  
XX  
XX Disclosure; Page 122-124; 175pp; English.  
XX  
XX The present invention provides the protein and coding sequences of  
XX monoterpene synthases from the grand fir. These include (-)-camphene  
XX synthase, (-)-beta-phellandrene synthase, terpinolene synthase, (-)-  
XX limonene/(-)-alpha-pinene synthase, limonene synthase, myrcene synthase  
XX and pinene synthase. The sequences can be used to produce transgenic  
XX plants expressing high levels of the enzymes, resulting in levels which  
XX are useful in protecting against and treating cancers, and to confer  
XX insect resistance on plants  
XX  
SQ Sequence 782 AA;

Query Match 51.6%; Score 2050.5; DB 4; Length 782;  
Best Local Similarity 50.8%; Pred. No. 2e-175;  
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;  
QY 6 GDGET--NPSAYDTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGECEIFLAYDRV 63  
DB 38 GDGESMITPSAYDTAWARVPAIDGSAARPQFPQTVDWILKNQLKDGSGWGIQSHFLSDRL 97  
QY 64 LNTLACLLTKIKNKGDIQVQKGVFVRKHEEMKDEAD-NHRPSGFEVVPFAMLDKAS 122  
DB 98 LATLSCLVLLKWNVGLQVEQIEFKSNLELVKDETDQDLSLVTDFEIFFPSLLREAO 157  
QY 123 LGLDLPYHLPFTISQIHQKQKLOKIPNLVNNHQTALLYSLGLEQDVVDWQEIITNLSR 182  
DB 158 LRGLPYDLPYIHLQTKQERLAKLSREIYAVPSPLLYSLEGIQDIVEWERIMEVQSQ 217  
QY 183 DGSFSLSPASTACVFMHTQNKCLHFLNVLKSGDVVPCYPLDLFERLWAVDTVERLG 242  
DB 218 DGSFSLSPASTACVFMHTGDAKCLEFLNSVMKEFGNVPCLYPVLLERLLIVDNI 277  
QY 243 IDRFYKKEIKESLDYVYRYWDAERGVGWACNPIPDVDDTAMGLRIILRHGYNVSSDYLE 302  
DB 278 IYRHFKEIKALDYVYRHNW-ERGIGWGRNLNADLETALGFRLLRLHRYNVSIPAIFD 336  
QY 303 NFRDEKGFDFCFAGQTOIGVTDNLNLYRCSQVCPGKIMEBEAKTFTTNHLQNALAKNA 362  
DB 337 NFKDANGKFCSTGQFNKDVASMLNLYRASQLAPPGENILDEAKSFATKYLREALEKSET 396  
QY 363 FDKWAVKDKLPGEVEYAIKYPWHRSMPLREARSYIEQFGSNDVMLGKTVYKMLYVSNKY 422  
DB 397 SSANNKQNLSEIKYALKTSWHSAPRVEAKRYCQVYRDPYARIKACVYKLPYVNNKFP 456  
QY 423 LELAKLDNMVQALHOKETQHIIVSWWRBESGFNDLFTTRQRPVEMYFVSVAVSMFEPEFAAC 482

457 LEIGKLDLFIQIHQEMKNVTSWFRDSDGLPFTFARERPLRFYFLVAAGTTEPQYAKC 516  
 483 RIAYAKTSCIAVLDDLDYDTHGSLDDDLKLFSEAVRRWDISVLDSVRDNLQKVCFLGLYNT 542  
 517 RFLPTKVAQLQTVLDDMDYDTYGLDELKLFTEAVRRWDLSTFENLDP-YMKLCVQIYDI 575  
 543 VNGFGKGLKEQGRDVLGYLRKWEGLLASYTEAEWSAAKYPTTNEYYVENAKVSIALA 602  
 576 VHEVAWEAEKEQGRDELVSFRKGMEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGQR 635  
 603 TVVLNSI-PFTGELLDPYILQQVDL--RSKFLHLVSLTGLINDTKTYOAEARNRGLVSS 659  
 636 ILLSGVLMDGQLLSQEALEKVDYPGRRVLTENLSRLADDTKTYRAKARGELASS 695  
 660 VOCYMRNPECTEERALSHPVGIIDNALKEINWELANPASNAFLCVRRLLFNTRARWOLF 719  
 696 IECYKMDHPECTEERALDHIYSILEPAVKELTREFLKP-DDVPFACKOMLFEETRVTMVI 754  
 720 YMYRDFGIGSDKEMKHVSRTLPDPV 745  
 755 FKDGDFGVSKLEVDHIKECLIEPL 780

## RESULT 8

ABB79392  
 ID ABB79392 standard; protein; 782 AA.  
 AC ABB79392;  
 DT 13-AUG-2002 (first entry)  
 XX Taxadiene synthase homologous protein sequence w85703.  
 DE Taxane; taxane synthesis; taxadiene synthase; cytostatic; anticancer;  
 KW enzyme.  
 XX Unidentified.  
 OS WO200240694-A2.  
 XX 23-MAY-2002.  
 PD 16-NOV-2001; 2001WO-DK000763.  
 PF 17-NOV-2000; 2000DK-00001730.  
 PR 29-NOV-2000; 2000US-0253843P.  
 XX (NOVO ) NOVOZYMES AS.  
 PA Vind J;  
 PI WPI; 2002-471624/50.  
 DR Heterologous production of taxane analogs related compound useful as  
 PT anticancer agent involves cloning a full-length taxane synthesis pathway  
 PT from a taxan-producing organism into a taxane-resistant host cell.  
 XX Example 3; Fig 1; 60pp; English.

The present invention describes the heterologous expression of taxane  
 analogues, comprising cloning a DNA sequence comprising a taxane  
 synthesis pathway, making a DNA construct in which the DNA sequence is  
 under control of regulatory elements, introducing the DNA construct into  
 a host cell, growing the host cell to produce taxane analogues, and  
 recovering the required taxane from the culture medium. Also described:  
 CC (1) an isolated DNA sequence comprising the taxane synthesis pathway;  
 CC (preferably taxol synthesis pathway); (2) an expression vector comprising  
 CC the isolated DNA sequence; and (3) a host cell comprising a taxane  
 CC synthesis pathway derived from a taxane-producing organism, plant or  
 CC tree, the taxane synthesis pathway is foreign to the host cell. The  
 CC taxane analogues have cytostatic activity, and so can be used as  
 CC anticancer agents. The present sequence represents a taxadiene synthase

CC homologous protein, which is given in an example from the present  
 CC invention  
 XX  
 SQ Sequence 782 AA;  
 Query Match 51.6%; Score 2050.5; DB 5; Length 782;  
 Best Local Similarity 50.8%; Pred. No. 2e-175;  
 Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;  
 QY 6 GDGET--NFSAYDTAWVARIPSIDGSGAQFPOTLQWILNNQLPDGSGWSECIFLAYDRV 63  
 DB 38 GDGESMITTSAYDTAWVARPAIDGSARQFPQPTVDWILKNQKDGSGWGQSHFLLSDRL 97  
 QY 64 LNTLACLLTLKWNKGDIOVKQGVFVRKHEMEMKDEAD-NHRPSPFVVFPAMLDRAKS 122  
 DB 98 LATLSCLVLLKWNVGDQVQEGIEPIKSNLELVKDETQDQSLVTFEITFPSSLREAQS 157  
 QY 123 LGLDLPHLPFTLSQIHQEKQKLOKIPLVNHHNQTALLYSLGLQDVVDVMOETNLQSR 182  
 DB 158 LRGLPYDLPYIHLQTKQERLAKLSREIYAVPSPLLYSLSGIQDIVEWERIMEVSQ 217  
 QY 183 DGSFLSSPASTACVFMHTQNKRCLEHFNFLVSLKFGDYVPCYHPLDLFERLWAVDTVERLG 242  
 DB 218 DGSFLSSPASTACVFMHTGDAKCLEFLNSVMIKFGNFVPCLYPVDLLERLLIVDNIVRLG 277  
 QY 243 IDRYFKGKIKESLDYVRYWDAERGVMARCNDIPDVEDTAMGLRILRLHGVNVSVDLE 302  
 DB 278 IYRHFEKEIKALDYVYRHWN-ERGIQWGLRNPADLETALGFRLRLHRYNVSIPAIFD 336  
 QY 303 NFRDEKDFPCFAGQTQIGVTDNLNLYRCOVCFPGCEKIMEERAKTFTTNHLQNALAKNA 362  
 DB 337 NFKDANGKFTCTGQFNKDVASMLNLYRASQAFAPGENILDEAKSFATKYLREALEKSET 396  
 QY 363 FDKWAVKDLPGVEYAIKYPWHRSMRLEARSYIEQFGSNDVWLKGTVMKLYVSNKY 422  
 DB 397 SSANNKQNLSEIKYALKTSWHAISVPRVEAREYCVYRPDYARIAKCVYKLPVNNKEF 456  
 QY 423 LELAKLDPMNQALHOKETQHIWVSWRSGFNDLTTTRQRPVEMYFSVAVSMPEPEPAAC 482  
 DB 457 LEIGKLDLFIQIHQEMKNVTSWFRDSDGLPFTFARERPLRFYFLVAAGTTEPQYAKC 516  
 QY 483 RIAYAKTSCIAVLDDLDYDTHGSLDDDLKLFSEAVRRWDISVLDSVRDNLQKVCFLGLYNT 542  
 DB 517 RFLPTKVAQLQTVLDDMDYDTYGLDELKLFTEAVRRWDLSTFENLDP-YMKLCVQIYDI 575  
 QY 543 VNGFGKGLKEQGRDVLGYLRKWEGLLASYTEAEWSAAKYPTTNEYYVENAKVSIALA 602  
 DB 576 VHEVAWEAEKEQGRDELVSFRKGMEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGQR 635  
 QY 603 TVVLNSI-PFTGELLDPYILQQVDL--RSKFLHLVSLTGLINDTKTYOAEARNRGLVSS 659  
 DB 636 ILLSGVLMDGQLLSQEALEKVDYPGRRVLTENLSRLADDTKTYRAKARGELASS 695  
 QY 660 VOCYMRNPECTEERALSHPVGIIDNALKEINWELANPASNAFLCVRRLLFNTRARWOLF 719  
 DB 696 IECYKMDHPECTEERALDHIYSILEPAVKELTREFLKP-DDVPFACKOMLFEETRVTMVI 754  
 QY 720 YMYRDFGIGSDKEMKHVSRTLPDPV 745  
 DB 755 FKDGDFGVSKLEVDHIKECLIEPL 780  
 RESULT 9  
 AAY06567  
 ID AAY06567 standard; protein; 817 AA.  
 XX AAY06567;  
 AC AAY06567;  
 XX 08-OCT-1999 (first entry)  
 DT B-alpha-bisabolene synthase computer-generated sequence.  
 DE B-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;  
 XX B-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;

gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.

KW	gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.	372	NEFDANGKFCSTGQFNKDVASMLNLYRASQLAPFGENILDEAKSFATKYLREALEKSET	431
OS	Synthetic.	363	FDKAVKQDLPGVEYAIKYPMHRSMPRLAEASYIEQFGSDVWMLGKTYYKMLYVSEKEY	422
OS	Abies grandis.	432	SSAMNNKQNLQSEIKYALKTSWHASVPRVEAKRYCQVYRPDYARIACVYKLPYVNNKF	491
XX	Key	423	LELAKLDENMVQALHOKETQHIVSWMRSEGFNDLFTTRQRPVEMYSFVAVSMPEPEAAC	482
FT	Peptide	492	LELGKLDNFNIIQSIHQEEMKNVTSWFRDGLPLFTFARERPLEFELVAAGTYEPOYAKC	551
FT	/note= "conserved element, Asp residues function in coordination of a divalent cation"	483	RIYAKTSCIVALLDDLYTHGSLDDLKLPSEAVRWMDISVLDSDVRDNLKVCFLGLYNT	542
XX	WO9937139-Al.	552	RLFTTKVACLQTVLDDMYDTYGTLDLKLFTFAVRWDLSFTENLDP- YMKLCYQIYYDI	610
XX	29-JUL-1999.	543	VNGFGKDLGKEQGRDVLGYLRKVGELGLASYSYKEAEWSAAKYVPTTFYVYENAKVSTALA	602
XX	21-JAN-1999;	611	VHEVAWEAEKEQGRDELVSFRKGWEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGOR	670
XX	22-JAN-1998;	603	TVVLSI-FFTGELLDPYILQQVDL--RSKFLHLVSLTGRINDTKYQAEKRGELVSS	659
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.	671	ILLSGVLIMDGLLQSEALEKVDYPRRVLTSLNLSLADDTTKYKAEKARGELASS	730
XX	Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;	660	VQCYMRNPECTEEREAALSHVYGIIDNALKELNLANPASNAPLCVRLLENTARVMOLF	719
XX	WPI; 1999-458569/38.	731	IECYMKDHPECTEEREAALDHIYSILEPAVKELTREFLKP-DDVPFACKKMLFEETRVTMWI	789
XX	N-PSDB; AAX87530.	720	YMYRDGFGISDKEMKOHVSRTLPDV 745	
XX	Grand fir (Abies grandis) sesquiterpene synthase genes and protein products, useful for sesquiterpenoid production.	790	FKDGDGFGVSKLEVKDHIKECLIEPL 815	
XX	Example 11; Page 98-101; 136pp; English.			
XX	This is the sequence of an E-alpha-bisabolene synthase polypeptide encoded by a computer-generated nucleic acid (see AAX87530) based on an isolated grand fir (Abies grandis) E-alpha-bisabolene synthase clone (see AAX87504). It contains conservative amino acid substitutions relative to the native sequence (see AAX06562). E-alpha-bisabolene synthase is a wound-inducible enzyme capable of generating multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl diphosphate. The invention provides gymnosperm sesquiterpene synthase nucleic acids and recombinant proteins, as well as expression vectors, and a method of imparting or enhancing production of a gymnosperm sesquiterpene synthase in a prokaryotic or eukaryotic host cell, especially in a plant, in order to impart, enhance or modify production of sesquiterpenoids, and hence to alter the flavour or fragrance of plant products, to improve disease resistance or to alter ecological interactions mediated by farnesyl diphosphate and its derivatives			
XX	Sequence 817 AA;			
XX	Query Match			
XX	Best Local Similarity 51.6%; Score 2050.5; DB 2; Length 817;			
XX	Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;			
QY	6 GDGET--NPSAYDTAWARIPSIDSGAPQFPQTQWILANNQIPDGSWGEICIFLAYDRV 63			
DB	73 GDGESMITPSAYDTAWARIPADISGAPQFPQTQWILANNQIPDGSWGEICIFLAYDRV 132			
QY	64 LNTLACLLTKTKWNGDIOVQKGVFVRKHEMDEAD-NHRPSGFVFPFAMLDKAS 122			
DB	133 LATLSVLLKRNWGLQVEQIEFIKNLELVKDETQDSLVITDFEIFFPSLLREAGS 192			
QY	123 LGLDLPYLPFISQIHQKQKQKLOKPLNVLNHNQTALLYSLLEGQDVVDWQBETHLQSR 182			
DB	193 LRGLPYDLPYIHLQTKQERLAKLSREIYAVSPSPLLYSLEGIQDIIVEMERIMEVQSQ 252			
QY	183 DGSFLSPASTACVFHMTQKRCILHFLNVLKFGDYVPCYPLDLFERLWAVDTVERLG 242			
DB	253 DGSFLSPASTACVFHMTQKRCILHFLNVLKFGDYVPCYPLDLFERLWAVDTVERLG 312			
QY	243 IDVEYKEIKESLDYVRYWDAERGVGWACNPIPDVDDTAMGLRLRLHGVNVSVDLE 302			
DB	313 IYRHFKEIKESLDYVRYWDAERGVGWACNPIPDVDDTAMGLRLRLHGVNVSVDLE 371			
QY	303 NFRDEKGDFFCFAGQTOIGVTDNLNLYRCQVCPGKEIMEEAKTFTTNLQNALAKNNA 362			

Db	372	NEFDANGKFCSTGQFNKDVASMLNLYRASQLAPFGENILDEAKSFATKYLREALEKSET	431
QY	363	FDKAVKQDLPGVEYAIKYPMHRSMPRLAEASYIEQFGSDVWMLGKTYYKMLYVSEKEY	422
Db	432	SSAMNNKQNLQSEIKYALKTSWHASVPRVEAKRYCQVYRPDYARIACVYKLPYVNNKF	491
QY	423	LELAKLDENMVQALHOKETQHIVSWMRSEGFNDLFTTRQRPVEMYSFVAVSMPEPEAAC	482
Db	492	LELGKLDNFNIIQSIHQEEMKNVTSWFRDGLPLFTFARERPLEFELVAAGTYEPOYAKC	551
QY	483	RIYAKTSCIVALLDDLYTHGSLDDLKLPSEAVRWMDISVLDSDVRDNLKVCFLGLYNT	542
Db	552	RLFTTKVACLQTVLDDMYDTYGTLDLKLFTFAVRWDLSFTENLDP- YMKLCYQIYYDI	610
QY	543	VNGFGKDLGKEQGRDVLGYLRKVGELGLASYSYKEAEWSAAKYVPTTFYVYENAKVSTALA	602
Db	611	VHEVAWEAEKEQGRDELVSFRKGWEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGOR	670
QY	603	TVVLSI-FFTGELLDPYILQQVDL--RSKFLHLVSLTGRINDTKYQAEKRGELVSS	659
Db	671	ILLSGVLIMDGLLQSEALEKVDYPRRVLTSLNLSLADDTTKYKAEKARGELASS	730
QY	660	VQCYMRNPECTEEREAALSHVYGIIDNALKELNLANPASNAPLCVRLLENTARVMOLF	719
Db	731	IECYMKDHPECTEEREAALDHIYSILEPAVKELTREFLKP-DDVPFACKKMLFEETRVTMWI	789
QY	720	YMYRDGFGISDKEMKOHVSRTLPDV 745	
Db	790	FKDGDGFGVSKLEVKDHIKECLIEPL 815	
RESULT 10			
AAY06566			
ID	AAY06566 standard; protein; 817 AA.		
XX	AAY06566;		
XX	08-OCT-1999 (first entry)		
XX	Grand fir E-alpha-bisabolene synthase.		
XX	E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;		
XX	gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.		
XX	Abies grandis.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 65	/note= "Ala in AG1 (AAY06563)"	
FT	Peptide	566..570	
FT	/note= "conserved element, Asp residues function in coordination of a divalent cation"		
XX	WO9937139-Al.		
XX	29-JUL-1999.		
XX	21-JAN-1999;	99WO-US001300.	
XX	22-JAN-1998;	98US-0072204P.	
XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.		
XX	Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;		
XX	WPI; 1999-458569/38.		
XX	N-PSDB; AAX87529.		
XX	Grand fir (Abies grandis) sesquiterpene synthase genes and protein products, useful for sesquiterpenoid production.		
XX	Example 11; Page 92-95; 136pp; English.		





anticancer agent involves cloning a full-length taxane synthesis pathway from a taxane-producing organism into a taxane-resistant host cell.

Example 3: Fig 1: 60pp; English.

The present invention describes the heterologous expression of taxane analogues, comprising cloning a DNA sequence comprising a taxane synthesis pathway, making a DNA construct in which the DNA sequence is under control of regulatory elements, introducing the DNA construct into a host cell, growing the host cell to produce taxane analogues, and recovering the required taxane from the culture medium. Also described: (1) an isolated DNA sequence comprising the taxane synthesis pathway, (preferably taxol synthesis pathway); (2) an expression vector comprising the isolated DNA sequence; and (3) a host cell comprising a taxane synthesis pathway derived from a taxane-producing organism, plant or tree, the taxane synthesis pathway is foreign to the host cell. The taxane analogues have cytostatic activity, and so can be used as anticancer agents. The present sequence represents a taxadiene synthase homologous protein, which is given in an example from the present invention

Sequence 817 AA:

51 68: score 2050.5; DB 5; Length 817;

every Match  
51.6%;  
score 2030:3, 22:3;  
pred. No. 2.2e-175;

7;	Gaps
9;	Indels
206;	Mismatches
379;	Conservative
152;	Indels
206;	Mismatches
50.8%;	Pred. NO. 2.28-1.75;
50.8%;	Pred. NO. 2.28-1.75;

[illegible]

RESIST 12

RESOL 12  
ABR79391

ABB/9391  
IN ABB79391 standard: protein; 817 AA.

ID  
YY  
ABB/5331

XX  
AC  
APB70391.

AC  
VV  
ABB/9391;

XX  
DT  
13-AUG-20DT  
13-AUG-20  
yy

5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528

DE Taxadiene

4  
5  
6  
7  
8  
9  
E

KW Taxane; t

KW enzyme.

XX 13-14

OS Unidentified

XX

PN WO2002406

XX

PD 23-MAY-20

XX

PF 16-NOV-20

XX

PR 17-NOV-20

Db	731	IECYMDHPECTEERLADHIYSILEPAVKELTREFLKP-DDVPACKKMLFEETRVTMTVI	789
Qy	720	YMYRDGFGISDKEMKHVSRTLPDPV	745
Db	790	FKDGDGFGVSKLEVKDHKECLIEPL	815
RESULT 13			
ABU09781	ID	ABU09781	standard; protein; 817 AA.
XX	AC	ABU09781;	
XX	AC	ABU09781;	
DT	06-AUG-2003	(first entry)	
XX	XX	White fir E-alpha-bisabolene synthase.	
XX	XX	Levopimaradiene synthase; ginkgolide; enzyme; Escherichia coli;	
KW	KW	Candida albicans; Kluyveromyces lactis; Saccharomyces; white fir;	
KW	KW	E-alpha-bisabolene synthase.	
XX	XX		
OS	OS	Abies grandis.	
XX	XX	US2002164736-A1.	
XX	XX	07-NOV-2002.	
XX	XX	07-JAN-2002; 2002US-00041007.	
XX	XX	05-JAN-2001; 2001US-0259881P.	
XX	XX	(MATSU) MATSUDA S P T.	
XX	XX	(SCHEP) SCHEPMANN H G.	
XX	XX	Matsuda SPT, Schepmann HG;	
XX	XX	WPI; 2003-352584/33.	
XX	XX	New purified and isolated nucleic acid sequence encoding a	
PT	PT	levopimaradiene synthase useful for producing pure ginkgolides in cells.	
XX	XX	Example 3; Fig 4; 37pp; English.	
XX	XX	The invention relates to a purified and isolated nucleic acid sequence	
CC	CC	encoding a levopimaradiene synthase. The nucleic acid is used for the	
CC	CC	preparation of a ginkgolide in a cell, an expression vector, a	
CC	CC	unicellular organism (e.g. Saccharomyces, Escherichia coli, Candida	
CC	CC	albicans or Kluyveromyces lactis), a yeast host cell, a plant	
CC	CC	(particularly Ginkgo biloba), a host cell, a transgenic plant or a seed	
CC	CC	of the transgenic plant, for the biosynthesis of ginkgolides and in	
CC	CC	molecular biology, molecular genetics, and organic chemistry. The nucleic	
CC	CC	acid is capable of producing large amounts of pure ginkgolides in a cost-	
CC	CC	effective manner, particularly in an organism capable of a high-yield	
CC	CC	ginkgolide-producing system. This sequence represents an E-alpha-	
CC	CC	bisabolene synthase polypeptide which shows similarity to the	
CC	CC	levopimaradiene synthase of the invention	
XX	XX	Sequence 817 AA;	
Query Match			
Best Local Similarity 51.6%; Score 2050.5; DB 6; Length 817;			
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;			
Qy	6	GDGET--NPSAYDTAWARIPSDGAPFPQTQWILNNQLPDGSGWGECEIFLAYDRV	53
Db	73	GDGSEMTTPSAYDTAWARIPADGSGAPFPQTQWILNNQLPDGSGWGECEIFLAYDRV	132
Qy	64	LNTLACLLTKIWNKGDIQVKGVEFKMEKMDAD-NHRPSGFVFPAMLEAKS	122
Db	133	LATLSCLVLLKNNVGLQVEQIEFKSNLELKDQDQSLVTDPEIIFPSLLREAS	192
Qy	123	LGDLPLPFIQIHOKRQKKQIKIPLNVLNNHQTALLYSLGLQDVVDWQBITNLOS	182

Db	193	LRGLPYDLPYTHLQTKQERLAKLSREIYAVPSPLLYSLGQIDIVEMERIMEVQSQ	255
Qy	183	DGSFLSSPASTACVFMHTQNKCLHFLNFVLKFGDYVPCHPVPLDLFLERLWAVDVTVERLG	243
Db	253	DGSFLSSPASTACVFMHTQNKCLHFLNFVLKFGDYVPCHPVPLDLFLERLWAVDVTVERLG	312
Qy	243	IDRYFKKIKESLDVYVRYWDAERGVMGWARCNPIPDVDDTAMGLRILRLHGVSSDVLK	302
Db	313	IYRHFEKEIKALDYVYRHN-ERGIWGRNLNPIADLETTLALGPRLLLRHRYNVSIPAIF	371
Qy	303	NPDDEKDFCFAGQTOIGVTDNLNLYCSQVCFPEKIMEBAKFTTNHLQNALAKNA	362
Db	372	NFKDANGKFCSTGQFNKDVASMLNLYRASQALPFGENILDEAKSFATKYLEALEKSET	431
Qy	363	FDKWAVKDLPGEVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLKGTYYKMLVYSNEKY	422
Db	432	SSANNKQNLQSIEIKYALKTSHWASVPRVEAKRYCQVYRDPYARIATKCVYKLPYVNNKFP	491
Qy	423	LELAKLDFNNVQALHOKETQHIVSWNRSGFNDLFTTQRQPVEMYFSVAVSMFEPEAAC	482
Db	492	LELGKLDNIIQSIHQEEMKNVTSNFRDGLPLFTFARERPLEFVFLVAAGTYEPOYAKC	551
Qy	483	RIAYAKTSCLAVIDLDLYDTHGSIDDLKLFSEAVRWDISVLDSVRDNLQKVCFLGLYNT	542
Db	552	RFULTKVAQLQTVLDDMDYDTGTLDELKLTFAVRRWDLSTFENLPD-YMKLCYQIYYDI	610
Qy	543	VNGFGKGLKEQGRDVLGYLRKRWEGLLASVTKBAEWSAAKVPTFNFYVENAKVSIALA	602
Db	611	VHEVAWEAKEQRELVSFRKGWEDYLLGYEEAEWLAAYVPTLDYIKNGITSIQOR	670
Qy	603	TVVLNSI-EFTGELLPPDYILQQVDL--RSKFLHLVSLTGRLLINDTKTYOABNRGELVSS	659
Db	671	ILLGVLIMDQQLLSQEALEKVDYVGRRLVLTSLNSLSRLADDTKTYKAEKARGELASS	730
Qy	660	VQCYMRENPCTEERLASHVYCIIDNALKELNWLANPASNAPLCVRRLLFTARVMQLF	719
Db	731	IECYMDHPECTEERLADHIYSILEPAVKELTREFLKP-DDVPACKKMLFEETRVTMTVI	789
Qy	720	YMYRDGFGISDKEMKHVSRTLPDPV	745
Db	790	FKDGDGFGVSKLEVKDHKECLIEPL	815
RESULT 14			
AAV06568	ID	AAV06568	standard; protein; 817 AA.
XX	AC	AAV06568;	
XX	AC	AAV06568;	
DT	08-OCT-1999	(first entry)	
XX	DE	E-alpha-bisabolene synthase computer-generated sequence.	
XX	KW	E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;	
XX	KW	gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.	
XX	OS	Synthetic.	
OS	OS	Abies grandis.	
XX	XX	Key	
XX	XX	Location/Qualifiers	
XX	XX	566..570	
XX	XX	/note= "conserved element, Asp residues function in	
XX	XX	coordination of a divalent cation"	
XX	XX	W09937139-A1.	
XX	XX	29-JUL-1999.	
XX	XX	21-JAN-1999; 99WO-US001300.	
XX	XX	22-JAN-1998; 98US-0072204P.	
XX	XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.	

Db	193	LRGLPYDLPIYHLLQTKROERLAKLSREIYAVPSPLLYSLSEIQIDIVERMERMEVQSQ	252
Qy	183	DGSFLSPASTACVFMHTQNKRLHFLNLFVLSKFGDYVPCPLDLPERLWAVDTVERLG	242
Db	253	DGSFLSPASTACVFMHTGDAKCLFLNLSVMKFGNFPVCLYPVDLLERLLVDNIVRLG	312
Qy	243	IDRYFKKEIKESLDYVRYWDAERGVCWACNIPDVEDDTAMGLRIILRHGYNVSSDYLE	302
Db	313	IYRHFKEIKELADYVYRHNW-ERGIWGLNPLADLETTALGFRLLRLHRYNVPAPFD	371
Qy	303	NFRDEKGFDFCFAGQTOIGVTNINLYRCQVCPGKIMEEAKTFTTNHLONAKNNA	362
Db	372	NFKDANGKFCSTGQFNKDVASMLNLYRASQLAPGENILDEAKSPATKYLREALESET	431
Qy	363	FDKAVKXKDLPGREYVAIKYPHRSMPRLREARSYIEQFGSNDVWMLGKTVYKMLYSNEKY	422
Db	432	SSAWNNKQLSQBIKVALKTSWHSVPRVEAKRYCQVYRDPYARIACVYKLPYNNKEP	491
Qy	423	LELAKLDFNMVQALHOKETQHIYVSWRESQFNDLTTRQRPVEMYSVAVSMPEPEAAC	482
Db	492	LELGLDFNIIQSIHQSEMKNVTSWFRDGLPLFTFARERPLRFYFLVAAGTYEPOYAKC	551
Qy	483	RIAYAKTSCLAVIDDLXDTGSLDDLKLPSEAVRRWDISVLDSVRDNLKVCFLGLYNT	542
Db	552	RFLFTKVALQTVDDMYDTYGLDELKLEFAVRWDLSTENLPD-YMKLCYQIYDI	610
Qy	543	VNGFGDKLKEQGRDVLGYLRKWEGLLASAKYTKAEWSAAKYPTFNVEYVENAKVSIALA	602
Db	611	VHEVAWEAEKEQGRVLSVPRKWDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGOR	670
Qy	603	TVVLNSI-FFTGELLDPYILQQVDL--RSKFLHLVSLTGRINDTKTYQABRNKELVSS	659
Db	671	ILLGLVIMDGLLSQEALEKVDYPCRRVLTSLNLSLRLADDTTKTYKAEKARGELASS	730
Qy	660	VOCYMEENPECTEERLADHIYSILEPAVKELTREFLKP-DDVPACKMLFEETRTMTVI	719
Db	731	IECYMDHPECTEERLADHIYSILEPAVKELTREFLKP-DDVPACKMLFEETRTMTVI	789
Qy	720	YMRDGFSGISKEMKHVSRTLPDPV	745
Db	790	FKDGDGFGVSKLEVKDHKECLIEPL	815
RESULT 14			
ID	AAV06568	standard; protein; 817 AA.	
XX	XX	AAV06568;	
XX	XX	08-OCT-1999 (first entry)	
DT	DT	E-alpha-bisabolene synthase computer-generated sequence.	
XX	XX	E-alpha-bisabolene synthase; grand fir; sesquiterpene synthase;	
XX	XX	gymnosperm; sesquiterpenoid; disease resistance; fragrance; flavour.	
XX	OS	Synthetic.	
OS	OS	Abies grandis.	
XX	XX	Key	
FT	FT	Location/Qualifiers	
FT	FT	566..570	
FT	FT	/note="conserved element, Asp residues function in	
FT	FT	coordination of a divalent cation"	
XX	XX	WO9937139-A1.	
XX	XX	29-JUL-1999.	
XX	XX	21-JAN-1999; 99WO-US001300.	
XX	XX	22-JAN-1998; 98US-0072204P.	
XX	XX	(UNIW ) UNIV WASHINGTON STATE RES FOUND.	

603 TVVLSI-FFTGELPDILOQVDL--BSKFLHLVSLTGRLLINDTKTYQAERNRGLVSS 659  
 671 ILLSGVLIMDQQLLSQALEKVDYPPGRVLTSLNLSRLADDTKTYKAERKARGELASS 730  
 660 VOCYMRNPCTEEREAALSHVYGIIDNALKELWELANPASNAPLCVRRLLFNARVMOLF 719  
 731 IECYMDHPCTEEREAALSHVYGIIDNALKELWELANPASNAPLCVRRLLFNARVMOLF 719  
 720 YMYRGGFISDKEMKOHVSRITLFDV 745  
 790 FKDGDFGVSKLEVXDHKEICIEPL 815

RESULT 15  
 AAW31655  
 ID AAW31655 standard; protein; 862 AA.  
 XX AAW31655;  
 XX 30-APR-1998 (first entry)  
 XX Pacific yew taxadiene synthase.  
 XX Pacific yew; taxadiene synthase; taxol biosynthesis; probe; immunoassay.  
 XX Taxus sp.  
 XX W09738571-A1.  
 XX 23-OCT-1997.  
 XX 15-APR-1997; 97WO-US006320.  
 XX 15-APR-1996; 96US-0015993P.  
 XX (UNIW) UNIV WASHINGTON STATE RES FOUND.  
 XX Wildung MR, Croteau RB;  
 XX WPI; 1997-526123/48.  
 XX N-PSDB; AAT97447.  
 XX DNA encoding Pacific yew taxadiene synthase - used to increase taxol biosynthesis.  
 XX Disclosure; Fig 2; 30pp; English.  
 XX This is a cloned Pacific yew taxadiene synthase (TS) protein. The sequence can be used to increase taxol biosynthesis into taxoids in transformed cells. The protein can be modified to improve efficiency, and to redirect the targeting of the expressed polypeptide. The modifications can be to alter catalytic efficiency, to modify substrate utilisation, and to alter product outcome. The antibodies specific for the mature Pacific yew TS polypeptide can be used to purify TS polypeptides, cloning TS homologues from Pacific yew or other plant species, and as probes for blots and immunoassays

Query Match 47.9%; Score 1901; DB 2; Length 862;  
 Best Local Similarity 47.5%; Pred. No. 7.2e-162; Indels 28; Gaps 9;  
 Matches 358; Conservative 149; Mismatches 219;

1 MFCSDGDTNPSAYTAWVARIPSI--DGSQAPQFPQTLOWLNNQLPDGSGGECIFL 58  
 126 MFVALGDGDISPAYTAWVARLATISSDSEKPRFPQALNWFNNQLQDGSWGIESHFS 185  
 59 AYDRVLTNLTCLTLKINWKGDIQVQKGVFVRKHEEMKDEADNHRPSPGVFVPAWLD 118  
 186 LCDRLNTTNSVIALSVKTKGSHVQVQAGAFIAENLRLNEE-DELSPD-FQIIFPALQ 243  
 119 EAKSLGLDLPYHLFFISQIHQKQKKL-----QKIPLNVLNHNHOTALLYSLGLQDV 171

XX PI Croteau RB, Crock JE, Bohlman J, Jetter R, Steele CL;  
 XX WPI; 1999-458569/38.  
 XX N-PSDB; AAX87531.  
 XX Grand fir (*Abies grandis*) sesquiterpene synthase genes and protein products, useful for sesquiterpenoid production.  
 XX Example 11; Page 104-106; 136pp; English.  
 XX This is the sequence of an E-alpha-bisabolene synthase polypeptide encoded by a computer-generated nucleic acid (see AAX87531) based on an isolated grand fir (*Abies grandis*) E-alpha-bisabolene synthase clone (see AAX87504). It contains conservative amino acid substitutions relative to the native sequence (see AAX06562). E-alpha-bisabolene synthase is a wound-inducible enzyme capable of generating multiple sesquiterpenes, especially E-alpha-bisabolene, from farnesyl diphosphate. The invention provides gymnosperm sesquiterpene synthase nucleic acids and recombinant proteins, as well as expression vectors, and a method of imparting or enhancing production of a gymnosperm sesquiterpene synthase in a prokaryotic or eukaryotic host cell, especially in a plant, in order to impart, enhance or modify production of sesquiterpenoids, and hence to alter the flavour or fragrance of plant products, to improve disease resistance or to alter ecological interactions mediated by farnesyl diphosphate and its derivatives

Query Match 51.6%; Score 2049.5; DB 2; Length 817;  
 Best Local Similarity 50.8%; Pred. No. 2.7e-175; Indels 9; Gaps 7;  
 Matches 379; Conservative 152; Mismatches 206;

6 GDGET--NPSAYDTAWVARIPSDGSAPOFPQTLOWLNNQLPDGSGGECIFLAVDRV 63  
 73 GDGSMITPSAYDTAWVARIPSDGSAPOFPQTLOWLNNQLPDGSGGECIFLAVDRV 132  
 64 LNTLACLTLLKINWKGDIQVQKGVFVRKHEEMKDEAD-NHRPSGFEVFPAMLEAKS 122  
 133 LATLSCLVLLKNNVGLDQVQKGVFVRKHEEMKDEAD-NHRPSGFEVFPAMLEAKS 192  
 123 LGLDLPYHLFFISQIHQKQKLOKIPLNVLNHNHOTALLYSLGLQDVVDMQEIITNLSR 182  
 193 LRLGLPDPYHLFFISQIHQKQKLOKIPLNVLNHNHOTALLYSLGLQDVVDMQEIITNLSR 252  
 183 DGSFLSSPASTACVFMHTQNKRLHFLNVLKFGDYVPCHPYPLDLFELNVAVDVTEBLG 242  
 253 DGSFLSSPASTACVFMHTQNKRLHFLNVLKFGDYVPCHPYPLDLFELNVAVDVTEBLG 312  
 243 IDRYPKKEIKESLDVYRWDARERGVGWACNPIPDVDDTAMGLRLRLHGYNVSVDVLE 302  
 313 IYRHFKEIKESLDVYRWDARERGVGWACNPIPDVDDTAMGLRLRLHGYNVSVDVLE 371  
 303 NFRDEKGFDFCFAGQIQIGVTNINLNYRCSQVCFPGKIMBEAKTPTTNHNLQNALAKNA 362  
 372 NFKDANGKFCISGQFNKDVASMLNLYRASQAFPGENILDEAKSPATKYLREALEKSET 431  
 363 FDKWAVKDLPGVEVYAIKFWHRSMPRLREARSYIEQFGSDNVLGLKTVYKMLYVNEKY 422  
 432 SSANWNNKQLSBIKVALKTSWASPRVBEAKRYCQVYRDPYARIKACVYKLPYVNEKF 491  
 423 LELAKLDNNVQALHOKETOHIVSWRESGFDNLTTRQRPVEMYSFVAVSMFEPFAAC 482  
 492 LEIGKLDNNVQALHOKETOHIVSWRESGFDNLTTRQRPVEMYSFVAVSMFEPFAAC 551  
 483 RIAYAKTSCIAVLDDLDYTHGSLDDKLSFSEAVRWDIISVLSVRDNLKVCFLGLYNT 542  
 552 RFLFTKVAQLQTVLDDMDYDTGTLDELKLFTEAVRRWDLSTFTNLPD-YMKLCYQIYDI 610  
 543 VNGFGKGLKEQGRDVLGKRLKWEGLASVYKAEWSAAKYVPTFNEYVENAKVSTALA 602  
 611 VHEVAWEAEKEQGRDVLGKRLKWEGLASVYKAEWSAAKYVPTFNEYVENAKVSTALA 670

Db 244 KAKALGINLPYDLPIFKYLTTRARLTDVSAADNI PANMLN-----ALEGLEEVI 295  
QY 172 DWOETINLOSROGSFLSSPASTACVPMHTONKRCLEHFLNVLSPKGDYVPCYHPDLDFER 231  
Db 296 DWNKIMRFOSKDGSLSSPASTACVLMNTGDEKCFTLNLLDKFGGCVPCMYSIDLLER 355  
QY 232 LWAVDTVERLGDIDRVPKKEIKESLDVYVYVWDAERGVCWARGCNPIPDVDDTAMGLILRL 291  
Db 356 LSLVNIIEHLGIRGPKQIKGALDIVYRHN-SERGIGMRDSLVDPDLNLTALGLTLRM 414  
QY 292 HGYNVSSDVLENFRDEKGGFFCFAGQTOIGVTDNLNLRYCSQVCFPGBEKIMBEAKTFTTN 351  
Db 415 HGYNVSSDVLENFRDEKGGFFCFAGQTHVELRSVNLFRASDLAPDERAMDDARKFAEP 474  
QY 352 HLONALAKNNAFDKAVAKDLPGEVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLKTV 411  
Db 475 YLREALA-----TKISTNTKLFKEIYVVEYEPWHSIIPREARSYIDSDDNYVMQKTL 529  
QY 412 YKMLVVSNEKYLELAKLDNMMQALHOKETOHTVSWWRESGFNDLFTTRQRPVEMYFSVA 471  
Db 530 YRMPLSNSKCLLELAKLDNIVQSLHQEBLKULTRWKESGMADINFTRHRVAEYFSSA 589  
QY 472 VSMFPEFAACRIAYAKTSLAVILDLDYDTHGSLDDKLFSBAVRWDISVLDSVRDQ 531  
Db 590 T--PEPEYSATRIAPTKICLQVLPDDMADI FATLDELKSFTEGVKXWDTSLHHEIPE-C 546  
QY 532 LKVCFLGLYNTVNGFGKGLKSGRDVLYGLRKWEGLLASYTKAEAWSAAKYVPTFNEY 591  
Db 647 MQTCPKWPKLMEEVNNDVVKQGRDMLAHIRKPMELYNFCYQEREWLEAGYIPTFEY 706  
QY 592 VENAKVSTALATVVLNSIPTGELLDPDYILOQVDLSKFLHLVSLTGRINDTKTYQAR 551  
Db 707 LKTYAISVGLGPCTLPQILLMGEVLKDDVVEKVHPSPNNFELVLSWRLTNDTKTYQAK 766  
QY 652 NRGEVSSVQCYMRENPECTEBEALSHVYGIIDNALKELNWEANPASNAPLCVRRLFN 711  
Db 767 ARGQOASGIACYMKNPGATEADAIXHICRVVDRALKEASFEYFKPSNDIPMGCKSPFN 326  
QY 712 TARVMQLFYMRDGFISDKEMKHVSRTLDPV 745  
Db 827 LRLCVQIFYKFDGYGIANEEIKDVKVYIDPI 860

Search completed: August 17, 2004, 21:00:13  
Job time : 49.5386 secs

THIS PAGE IS BLANK

Result No.	Score	Match	Query	ID			Description
				Length	DB	ID	
1	2735.5	68.9	868	4	US-09-398-395A-56	Sequence 56, Appl	
2	2735.5	68.9	868	4	US-09-887-586A-56	Sequence 56, Appl	
3	2735.5	68.9	868	4	US-09-895-752-56	Sequence 56, Appl	
4	2735.5	68.9	868	4	US-09-903-012B-56	Sequence 56, Appl	
5	2735.5	68.9	868	4	US-09-900-797-56	Sequence 56, Appl	
6	2050.5	51.6	782	4	US-09-360-545-16	Sequence 16, Appl	
7	2050.5	51.6	782	4	US-09-398-395A-46	Sequence 46, Appl	
8	2050.5	51.6	782	4	US-09-887-586A-46	Sequence 46, Appl	
9	2050.5	51.6	782	4	US-09-895-752-46	Sequence 46, Appl	
10	2050.5	51.6	782	4	US-09-903-012B-46	Sequence 46, Appl	
11	2050.5	51.6	782	4	US-09-900-797-46	Sequence 46, Appl	
12	2050.5	51.6	817	3	US-09-234-393-38	Sequence 13, Appl	
13	2050.5	51.6	817	3	US-09-234-393-38	Sequence 38, Appl	
14	2050.5	51.6	817	3	US-09-234-393-40	Sequence 40, Appl	
15	2050.5	51.6	817	4	US-09-865-171-13	Sequence 13, Appl	
16	2050.5	51.6	817	4	US-09-865-171-38	Sequence 38, Appl	
17	2050.5	51.6	817	4	US-09-865-171-40	Sequence 40, Appl	
18	2049.5	51.6	817	3	US-09-234-393-42	Sequence 42, Appl	
19	2049.5	51.6	817	4	US-09-865-171-42	Sequence 42, Appl	
20	1901	47.9	862	3	US-09-315-861-2	Sequence 2, Appl	
21	1901	47.9	862	4	US-09-398-395A-44	Sequence 44, Appl	
22	1901	47.9	862	4	US-09-887-586A-44	Sequence 44, Appl	
23	1901	47.9	862	4	US-09-895-752-44	Sequence 44, Appl	
24	1901	47.9	862	4	US-09-903-012B-44	Sequence 44, Appl	
25	1901	47.9	862	4	US-09-900-797-44	Sequence 44, Appl	
26	1901	47.9	862	4	US-09-593-253-2	Sequence 2, Appl	
27	1381.5	34.8	811	4	US-09-614-913-22	Sequence 22, Appl	

121 KSLGLDLPYHLPFIQIHOQRQKQKQIPLNVLNHNHTALLYSLEGLQDVVDVHQBETNLQ 180  
246 KILGLDLPDLPFLKQIIEKREAKLARIPTDVLALPTTLLYSLEGLQETVDQKIMKLQ 305  
181 SRDGSFLSPASTACVFMHTQNRCLHFLNVLKSGDYVPCYPLDLPFLWADIVR 240  
306 SKDGSFLSPASTAAVFMRTGNKCLDFLNFVLFKFGNHVPCYPLDLPFLWADIVR 365  
241 LGTDVFKKEIKESLDVYRYWDAERGVGWARCNPIPDVDDTAMGLRILHLHGYNVSSDV 300  
366 LGTDVFKKEIKESLDVYRYWDAERGVGWARCNPIPDVDDTAMGLRILHLHGYNVSSDV 424  
301 LENFRDEKGFAGQTOIGVTDNLNLYRCQVCPGKIMEEAKFTTNHLONALAKN 360  
425 LKTFRDENGEGFCFLGQTOGRVTDMLNVRCSHVSPGETIMEEAKLCTERYLNALENV 484  
361 NAFDKWAKKDLPGVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLKTYKMLYVSN 420  
485 DAFDKWAKKNIIRGEVEYALKYPMHKSMPLEARSYIENTGPDVWLKTYKMLYVSN 544  
421 KYLELAKLDFNMVQALHOKETQIHVSWRESGNDLFTTRQPVMEYFVSVAVSMPEPEFA 480  
545 KYLELAKLDFNMVQALHOKETQIHVSWRESGNDLFTTRQPVMEYFVSVAVSMPEPEFA 604  
481 ACRIAYAKTSCLAVILDDLDYTHGSLDLDLKFSAVRWDISVLDSDVRDNOLKVCFLGLY 540  
605 KCREVYTKSNFTVILDDLDYTHGSLDLDLKFSAVRWDISVLDSDVRDNOLKVCFLGLY 663  
541 NTVNGFGKGLKEQGRDVLGYLRKRWEGLLASVYKAEWSAAKVPPTFNEVVENAKVIA 600  
664 NTFNDIAKEGRQGRDVLGYLRKRWEGLLASVYKAEWSAAKVPPTFNEVVENAKVIA 723  
601 LATVVLNSIPTGBELLDPYILQOVDLRSKFLHLVSLTGRILNDTKTYQAEARNRGELVSV 660  
724 LGTVVLISALFTGEVLTDVLSKIDRESRFLQMLGLTGRVNDTKTYQAEARNRGELVSV 783  
661 QCYMRENPCTEERALSHPYGIIDNALKELNWLANPASNAPLCVRRLLFNFTARVMQLFY 720  
784 QCYMRENPCTEERALSHPYGIIDNALKELNWLANPASNAPLCVRRLLFNFTARVMQLFY 841  
721 MYRQDGFIS-DKEMKDHVSRTLFDPA 746  
842 MQDGLTSLHDMIEKHEVKNCLFQPA 868

301 LENFRDEKGFAGQTOIGVTDNLNLYRCQVCPGKIMEEAKFTTNHLONALAKN 360  
425 LKTFRDENGEGFCFLGQTOGRVTDMLNVRCSHVSPGETIMEEAKLCTERYLNALENV 484  
361 NAFDKWAKKDLPGVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLKTYKMLYVSN 420  
485 DAFDKWAKKNIIRGEVEYALKYPMHKSMPLEARSYIENTGPDVWLKTYKMLYVSN 544  
421 KYLELAKLDFNMVQALHOKETQIHVSWRESGNDLFTTRQPVMEYFVSVAVSMPEPEFA 480  
545 KYLELAKLDFNMVQALHOKETQIHVSWRESGNDLFTTRQPVMEYFVSVAVSMPEPEFA 604  
481 ACRIAYAKTSCLAVILDDLDYTHGSLDLDLKFSAVRWDISVLDSDVRDNOLKVCFLGLY 540  
605 KCREVYTKSNFTVILDDLDYTHGSLDLDLKFSAVRWDISVLDSDVRDNOLKVCFLGLY 663  
541 NTVNGFGKGLKEQGRDVLGYLRKRWEGLLASVYKAEWSAAKVPPTFNEVVENAKVIA 600  
664 NTFNDIAKEGRQGRDVLGYLRKRWEGLLASVYKAEWSAAKVPPTFNEVVENAKVIA 723  
601 LATVVLNSIPTGBELLDPYILQOVDLRSKFLHLVSLTGRILNDTKTYQAEARNRGELVSV 660  
724 LGTVVLISALFTGEVLTDVLSKIDRESRFLQMLGLTGRVNDTKTYQAEARNRGELVSV 783  
661 QCYMRENPCTEERALSHPYGIIDNALKELNWLANPASNAPLCVRRLLFNFTARVMQLFY 720  
784 QCYMRENPCTEERALSHPYGIIDNALKELNWLANPASNAPLCVRRLLFNFTARVMQLFY 841  
721 MYRQDGFIS-DKEMKDHVSRTLFDPA 746  
842 MQDGLTSLHDMIEKHEVKNCLFQPA 868

RESULT 3  
US-09-895-752-56  
; Sequence 56, Application US/09895752  
; Patent No. 6559297  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6559297, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/895,752  
; CURRENT FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: 09/398,395  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 56  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Abies grandis  
US-09-895-752-56

Query Match 68.9%; Score 2735.5; DB 4; Length 868;  
Best Local Similarity 67.6%; Pred. No. 2e-270;  
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;  
1 MFQSGMDGTPNPSAYDTAWARIPSDGAPQPTQLWLNQPLDGSWGEEICFLAY 60  
126 MFRCMGYGETNPSAYDTAWARIPSDGAPQPTQLWLNQPLDGSWGEEICFLAY 185  
61 DVLNLTACLLTLTKMKNKGDIVQKGVFVRKIMEEMKDEADNHRPSGFVFPAMLDEA 120  
186 DRILATLACLLTLTKMKNKGDIVQKGVFVRKIMEEMKDEADNHRPSGFVFPAMLDEA 245





Db	784	QCTMKDHPKISEBEALQHVYVMENALELNREFVN--NKIPDIYKLVFETARIMOLFY	841
Qy	721	MYRDGFGIS-DKEMKDHVSRTLPDPA	746
Db	842	MOGDGLTSLHDMIBEKHKVKNCLFPQVA	868
RESULT 6			
US-09-360-545-16			
; Sequence 16, Application US/09360545			
; Patent No. 6429014			
; GENERAL INFORMATION:			
; APPLICANT: Croteau, Rodney B			
; APPLICANT: Bohlmann, Jorg			
; APPLICANT: Steele, Christopher L			
; APPLICANT: Phillips, Michael A			
; TITLE OF INVENTION: MONOTERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)			
; FILE REFERENCE: wsur13885			
; CURRENT APPLICATION NUMBER: US/09/360,545			
; CURRENT FILING DATE: 1999-07-26			
; EARLIER APPLICATION NUMBER: 60/052,249			
; EARLIER FILING DATE: 1997-11-07			
; EARLIER APPLICATION NUMBER: PCT/US98/14528			
; EARLIER FILING DATE: 1998-07-10			
; NUMBER OF SEQ ID NOS: 107			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 16			
; LENGTH: 782			
; TYPE: PRT			
; ORGANISM: Abies grandis			
US-09-360-545-16			
Query Match 51.6%; Score 2050.5; DB 4; Length 782;			
Best Local Similarity 50.8%; Pred. No. 2.2e-200;			
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;			
Qy	6	GDGET--NPSAYDTAWARIPIISIDSGAPQPOTLQWILNNQLPDGSGEGECIFLAYDRV	63
Db	38	GDGESMITPSSAYDTAWARVPAIDSGARPPQPTVVDWILKQKDKGSGWGIQSHFLLSDEL	97
Qy	64	LNTLACLTATLTKWNGDIQVQGVFVRKHEEMKDEAD-NHRPSGFEVVPFAMIDEAKS	122
Db	98	LATLSCLVLLKWNVGDLOVEGGIEFIKSNLELVKDETDQDSLVTDDEIFIPFSLREAGS	157
Qy	123	LGLDLPYHLPFTSQHQRKQKLOKIPNLVNHHTALLYSLEGLOVDVDMQEIITNLQSR	182
Db	158	LRLGLPYDLPYTHLQTKRQERLAKLSREIYAVFSPLLYSLEGIOIIVEMERIMEVQSQ	217
Qy	183	DGSFLLSSPASTACVFMHTQKRECLHFLNVLKFGDYVPCHPYDLPDLFERLWAVDTVERLG	242
Db	218	DGSFLLSSPASTACVFMHTGDAKLEFLNSVMIKFGNFVCLYPVDLLERLLIVDNIVRLG	277
Qy	243	IDRYFKKEIKESLDVYRYWDAERGVMARCNPIDVDVDTAMGLRLRLHGHVNVSSDVL	302
Db	278	IYHFEKEIKEALDVYRHN-ERGIGWGLNPIADLETTALGFRLLRLHRYNVSPAIFD	336
Qy	303	NFRDSKGFDFCFAGQTOIGVTNLIYRCQVCFPEKIMEEAKTFTTNHQLNALAKNA	362
Db	337	NFKDANGFKICTGQFNKDVASMLNLYRASQALFPGENILDEAKSFATKYLREALEKSET	396
Qy	363	FDKWA VKDLPGEVEYAIKYPWHRSMRLEARSYIEQFGSNDVWLGKTVYKMLYVSNKY	422
Db	397	SSAANNKQNSQEIKAITSWHAISVPRVEAKRYCQYRDPDYARIKAVVYKLPVYVNEKF	456
Qy	423	LELAKLDFNMVQALHOKETQHVSVWRESGNDLTTTRQRPVEMYSVAVSMPEPEAAC	482
Db	457	LELGKLDNFIIQSIHOEEMKNVTSWFRDGLPLFTFARERPLFFYFLVAAGTPEPQAKC	516
Qy	483	RIAYAKTSCLAVLTDLDYTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLYNT	542
Db	517	RFUFTKACVACVQVTDVDDMDYDTYGLDEBKLFTEAVRRWDLSTENLPD-YMKLCVQIYDI	575
Qy	543	VNGFGKDGKLEQGRDVLGYLRKYWEGLLASYTEKAEWSAAKYVPTTFNEYENAKVSTALA	602



38 GDGESMITPAYSADTAWARVPAIDGSRPQFPQTVDMILKNQKOGSWGIOHFLSDRL 97  
64 INTLACLLTKWNKGDIOVOKGVFVRKHEMDEAD-NHRPSGFVFPVPAWLDEAKS 122  
98 LATLSCVLLKNVNGDLQVQEQIEFKSNLELVKDBTDQDLSLVTDFEIFFSLLREAOQS 157  
123 LGIDLPHLPFISQIHQKQKLOKIPNLVNLHNOFTALLYSLGLQDVVDWQOETINLQSR 182  
158 LRLGLPYDLPYIHLQTKQERLAKLSREBIYAVPSPLYSLEGIQDIVEMERIMEVQSO 217  
183 DGSFLSPASTACVFMHTQNKRCLEHFNFLVLSKFGDVPCHPDLDFERLWAVDVTBRLG 242  
218 DGSFLSPASTACVFMHTGDKLEFLNVMIRKFGNVPCLPYVDLLERLLIVDNIVRLG 277  
243 IDRYEKEIKESIDYVYVYMDAERGVCWRCNIPDVDDTAMGLRIILRLHGVNVSVDVLE 302  
278 IYRHFKEIKESIDYVYVYMDAERGVCWRCNIPDVDDTAMGLRIILRLHGVNVSVDVLE 336  
303 NFRDEKGFPCFAGQTOIGVTDNLNLYRCSQVCPGKIMEEAKTFTTNHLONALAKNNA 362  
337 NFKDANGKFCSTGQFNKDVASMLNLYRASQLAPPGENILDEAKSFATKYUREALEKSET 396  
363 FDKWAVKDLPGVEVYAIKYPWHRSMRLEARSYIEQFSGNDVWLGKTVYKMLYVSNKY 422  
397 SSANNKONLSQEKIYALKTSWHSASVPRVEAKRYCQYRDPYARIACVYKLPYVNEKF 456  
423 LELAKLDFNMVQALHOKETQIHVSWRESGNDFLTFTRQRPVEMVFSVAVSMFEPEFAAC 482  
457 LELKGLDFNIIQSHOEBEMKNTSWFRDGLPLFTFARERPLEFYFLVAAGTYEPOYAKC 516  
483 RIAYAKTSLAVILDDLYDTHGSLDDKLSEAVRRWDISVLDSVRDNOLKVCFLGLYNT 542  
517 RFLFTKVAQLQTVLDDMYDTYGTDELKLFTEAVRRWDLSTFNLPD-YMKLCYIYYDI 575  
543 VNGFGKDLKEQGRDVLGYLRKWNGLASYTEAEWSAAKYVPTFNEYVENAKVSIALLA 602  
576 VHEVAWEAKEQGRRELVSFFKRWEDYLLGYEEAEWLAAYVPTLDEYIKNGITSIGOR 635  
603 TVVLNSI-FFTGELLDPYILOQVDL--RSKFLHLVSLTGRINDTKTYOABNRGELVSS 659  
636 ILLSGVLIMDQQLSQALEKVDYPRRRVLTSLNLSIRLADDTKYAKKARGELASS 695  
660 VOCYMRNPECTEERALSHTVGIIONALKELNWLANPASNAPLCVRRLLFNTRVMOQLF 719  
696 IECYMKOHPECTEERALSHTVGIIONALKELNWLANPASNAPLCVRRLLFNTRVMOQLF 754  
720 YMYRDFGIGSDKEMKHVSRTLFDPV 745  
755 FKDGDFGVSKLEVKDHKECLIEPL 780

RESULT 10  
US-09-903-012B-46  
; Sequence 46, Application US/09903012B  
; Patent No. 6569656  
; GENERAL INFORMATION:  
; APPLICANT: Chappell, Joseph  
; APPLICANT: No. 6569656, Joseph P.  
; APPLICANT: Starks, Courtney M.  
; APPLICANT: Manna, Kathleen R.  
; TITLE OF INVENTION: SYNTHASES  
; FILE REFERENCE: 07678-025001  
; CURRENT APPLICATION NUMBER: US/09/903,012B  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/100,993  
; PRIOR FILING DATE: 1998-09-18  
; PRIOR APPLICATION NUMBER: 60/130,628  
; PRIOR FILING DATE: 1999-04-22  
; PRIOR APPLICATION NUMBER: 60/150,262  
; PRIOR FILING DATE: 1999-08-23  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 46  
; TYPE: PRT  
; ORGANISM: Abies grandis  
US-09-895-752-46

Query Match 51.6%; Score 2050.5; DB 4; Length 782;  
Best Local Similarity 50.8%; Pred. No. 2.2e-200;  
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;  
6 GDGET--NPASAYDTAWARVPAIDGSRPQFPQTVDMILKNQKOGSWGIOHFLSDRL 63

```
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-903-012B-46

Query Match      51.6%; Score 2050.5; DB 4; Length 782;
Best Local Similarity 50.8%; Pred. No. 2.2e-200;
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;

QY 6 GGET--NPSAYDTAWARIPSDGSAFPQTQLOWILNNQLPDGSGWGBECIFLAYDRV 63
Db 38 GDGESMITPSAYDTAWARVPAIDGSARPOPTVDWILKNQLKDGSGWGIQSHFLSDRL 97
QY 64 IANTLACLLTLKWNKCDIQVQKGVFVRKHEMKDEAD-NHRPSGFEVVPFAMLEAKS 122
Db 98 LATLSCVLLKWNVGLQVEQIEFIKSNLELVKDETDQDSLVTPDFEIIFFSLLREAS 157
QY 123 LGLDLPHYLPFTSOIHOKRQKLOKIPNLVHNHOTALLYSLGLQDVVDWQBITNQSR 182
Db 158 LRGLPYDLPYHLTKQERLAKLSREIYAVPSPLLYSLGIIQDIWEVERIMEVQSQ 217
QY 183 DGSFLSSPASTACVFMHTQNKRCLEHFNLFVLSKFGDYVPCHPYDLDFERLMAVDTVBERLG 242
Db 218 DGSFLSSPASTACVFMHTGDAKCLEFLNSWMIKFGNFVPCLYPVDLLERLLIVDNIVRLG 277
QY 243 IDRYFKKEIKESLDYVRYWDAERGVGWACRNPIPDVDDTAMGLRLILRHGYNVSSDVLE 302
Db 278 IYRHFKEIKALDYVYRHNW-ERGIWGLRNPIDLETALGFRLRLRLHRYNVSPAFID 336
QY 303 NFRDEKGFDFCAGQTOIGVTDNILYRCQVCPGCKIMEEAKTFTTNHLQNALAKNA 362
Db 337 NFKDANGKFCSTGQFNKDVASMLNLYRASQLAFPGENILDEAKSPATKYLRLEAKSET 396
QY 363 FDKWAVKDLPGVEYAIKYPHRSMPRLARSYIEQFGSNDVWLKTYVYKMLYVSNKY 422
Db 397 SSANNKQNLQSEIKYALKTSWHSVPRVEAKRYCQVYRDPYARIKACVYKLPYVNNKEF 456
QY 423 LELAKLDFNMVQALHOKETQHIVSWRRESGFNDLTTTRQRPVEMYSVAVSMPEPEFAAC 482
Db 457 LELGKLDNFIIQSIHOEEMKNVTSWFRDGLPLFTFARERPLEFYLVAAGTYEPQYAKC 516
QY 483 RIAYAKTSCLAVIDLDLYTHGSLDLDLKLFSAVRRWDISVLDSVRDNLKVCFLGYNT 542
Db 517 RELFTKVACLQTVLDDMDYITGLDLSKLFTFARVRWDLSFTENLPD-YMKLCQIYDI 575
QY 543 VNGFGKDLGKQGRDVLGYLRKVEGLLASYTKAEASAAKYVPTFNEVYENAKVSIALA 602
Db 576 VHEVAWEAKEQGRRELVSFRKGMEDYLLGYEEAEMLAAEYVPTLDEYIKNGITSIGOR 635
QY 603 TVVLNSI-PFTGELLDPYILQVLDL--RSKFLHLVSLTGRLLINDTKTYQABNRNGBLVSS 659
Db 636 ILLLSGLVIMDGLLSQEALEKVDYVPGRRVLTSLNSLISRLADDTTKTYKAEKARGELASS 695
QY 660 VOCYRENPCTEERALSHTVGIIDNALKELNLANPASNAPLCVRRLLFNTARVWOLF 719
Db 696 IECYMKDHPECTEERALSHTVGIIDNALKELNLANPASNAPLCVRRLLFNTARVWOLF 754
QY 720 YMYRDFGIGSDKEMKHVSRTLPDPV 745
Db 755 FKDGDFGVSKLEVKHKECLIEPL 780

RESULT 11
US-09-900-797-46
; Sequence 46, Application US/09900797
; Patent No. 6645762
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. 66457621, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
```

Thu Aug 26 08:27:55 2004

```

660 VCYMEENPECTEERALSHVYGIIDNALKELNLANPASNAFLCVRRLLENFNTARNWOLF 719
731 IECYMKDHPECTEEREAALDHIYSILEPAVKELTEFLKP-DDVPFACKKMLFEETRVTWVI 789
720 YMYRDGFGISDKEMKDHVSRTLPDPV 745
790 FKDDGDFGVSKLEVKDHIKECLIEPL 815

RESULT 13
US-09-234-393-38
; Sequence 38, Application US/09234393A
; Patent No. 6265639
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Crock, John E
; APPLICANT: Bohlman, Jorg
; APPLICANT: Jetter, Reinhard
; APPLICANT: Steele, Christopher L
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
; FILE REFERENCE: WSUR113345
; CURRENT APPLICATION NUMBER: US/09/234,393A
; EARLIER FILING DATE: 1999-01-20
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-234-393-13

Query Match 51.6%; Score 2050.5; DB 3; Length 817;
Best Local Similarity 50.8%; Pred. No. 2.4e-200;
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;

QY 6 GDGET--NPSAYDTAWARIPISIDSGAPQPPQTLOWILNNOLPDGSGWGEICIFLAYDRV 63
DB 73 GDGESMTPTSAVDATAWARVPAIDGSAAPQPPQTVDMLKNQKDGSGWIGSHFLLSRL 132
64 LNTLACLLTLKWNKGDIOVQGVFVRKHEEMKDEAD-NHRPSGFVFPVPMALDEAKS 122
133 LATLSCLVLLKWNVDLQVEGIEFIKSNLELVKDETDQSLVTDFFIIFPSLLREAQS 192
123 LGLDLPYHLPTFSIQHQKQKLOKIPNLVNLHHTALLYSLEGLQDVVDWQEIITNLSR 182
193 LRLGLPYDLPYHLPTFSIQHQKQKLOKIPNLVNLHHTALLYSLEGLQDVVDWQEIITNLSR 252
183 DGSFLSSPASTACVPMHTQNKCHFLNPLVLSKFGDYVPCHYPDLDFERLWAVDVTVERLG 242
253 DGSFLSSPASTACVPMHTQNKCHFLNPLVLSKFGDYVPCHYPDLDFERLWAVDVTVERLG 312
243 IDRYEKEIKESLDVYRYWDAERGVMGWARCNPIPDVDDTAMGLRLRLHLYNVSSDVL 302
313 IYRHEKEIKEALDYYRHW-N-ERGIGWGRNLPIADLETTALGFRLLRLHLYNVSPAIFD 371
303 NFRDEKGFPCFAGQTOIGVTNLIYRCSQVCFPEKIMEBAKTFTTNHLQNALAKNA 362
372 NFKDANGKFCSTGQFNKDVASMLNLYRASQAFPGENILDEAKSFATKYLREALEKSET 431
363 FOKWAVKDLPGVEVEAIKYPWHRSMPLREARSYIEQFGSDNVLMGTVMYKMLYVSNKY 422
432 SSANWKNQLSQEIYKALKTSHASVPRVEAKRYCOVYRDPYARIKACVYKLPYVNNKEF 491
423 LEIADLDFNNQALHOKETQHIVSWWRSEGFNDLTFTROPVEMVYFSAVSMPEPEFAAC 482
492 LEIADLDFNNQALHOKETQHIVSWWRSEGFNDLTFTROPVEMVYFSAVSMPEPEFAAC 551
483 RIAYAKTSCLAVLDDLYDTHGSLDLDLKFSEAVRWDISVLDSVRDNLKVCFLGLYNT 542
552 RFLFTKVAQLQTVLDDMDYTYGTLDLDELKLFTEAVRWDLSFTENLPD-YMKLCYQIYYDI 610
543 VNGFGKGLKEQGRDVLGYLRKWKVEGLASYTEAEWSAAKYPTTNEVYENAKVIALA 602
611 VHEVAEAEKEQGRDVLGYLRKWKVEGLASYTEAEWSAAKYPTTNEVYENAKVIALA 670
603 TVVLNSI-PFTGELLDPYILQOVDL--RSKFLHVLSTGLINDTKTYOAEARNRGLVSS 659
671 ILLSGLVIMDQGLLSQEALEKVDYPGRRVLTENLSLSRLADDTKYKAEKARGELASS 730

```

```

660 VCYMEENPECTEERALSHVYGIIDNALKELNLANPASNAFLCVRRLLENFNTARNWOLF 719
731 IECYMKDHPECTEEREAALDHIYSILEPAVKELTEFLKP-DDVPFACKKMLFEETRVTWVI 789
720 YMYRDGFGISDKEMKDHVSRTLPDPV 745
790 FKDDGDFGVSKLEVKDHIKECLIEPL 815

RESULT 13
US-09-234-393-38
; Sequence 38, Application US/09234393A
; Patent No. 6265639
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Crock, John E
; APPLICANT: Bohlman, Jorg
; APPLICANT: Jetter, Reinhard
; APPLICANT: Steele, Christopher L
; TITLE OF INVENTION: SESQUITERPENE SYNTHASES FROM GRAND FIR (ABIES GRANDIS)
; FILE REFERENCE: WSUR113345
; CURRENT APPLICATION NUMBER: US/09/234,393A
; EARLIER FILING DATE: 1999-01-20
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 817
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-234-393-38

Query Match 51.6%; Score 2050.5; DB 3; Length 817;
Best Local Similarity 50.8%; Pred. No. 2.4e-200;
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;

QY 6 GDGET--NPSAYDTAWARIPISIDSGAPQPPQTLOWILNNOLPDGSGWGEICIFLAYDRV 63
DB 73 GDGESMTPTSAVDATAWARVPAIDGSAAPQPPQTVDMLKNQKDGSGWIGSHFLLSRL 132
64 LNTLACLLTLKWNKGDIOVQGVFVRKHEEMKDEAD-NHRPSGFVFPVPMALDEAKS 122
133 LATLSCLVLLKWNVDLQVEGIEFIKSNLELVKDETDQSLVTDFFIIFPSLLREAQS 192
123 LGLDLPYHLPTFSIQHQKQKLOKIPNLVNLHHTALLYSLEGLQDVVDWQEIITNLSR 182
193 LRLGLPYDLPYHLPTFSIQHQKQKLOKIPNLVNLHHTALLYSLEGLQDVVDWQEIITNLSR 252
183 DGSFLSSPASTACVPMHTQNKCHFLNPLVLSKFGDYVPCHYPDLDFERLWAVDVTVERLG 242
253 DGSFLSSPASTACVPMHTQNKCHFLNPLVLSKFGDYVPCHYPDLDFERLWAVDVTVERLG 312
243 IDRYEKEIKESLDVYRYWDAERGVMGWARCNPIPDVDDTAMGLRLRLHLYNVSSDVL 302
313 IYRHEKEIKEALDYYRHW-N-ERGIGWGRNLPIADLETTALGFRLLRLHLYNVSPAIFD 371
303 NFRDEKGFPCFAGQTOIGVTNLIYRCSQVCFPEKIMEBAKTFTTNHLQNALAKNA 362
372 NFKDANGKFCSTGQFNKDVASMLNLYRASQAFPGENILDEAKSFATKYLREALEKSET 431
363 FOKWAVKDLPGVEVEAIKYPWHRSMPLREARSYIEQFGSDNVLMGTVMYKMLYVSNKY 422
432 SSANWKNQLSQEIYKALKTSHASVPRVEAKRYCOVYRDPYARIKACVYKLPYVNNKEF 491
423 LEIADLDFNNQALHOKETQHIVSWWRSEGFNDLTFTROPVEMVYFSAVSMPEPEFAAC 482
492 LEIADLDFNNQALHOKETQHIVSWWRSEGFNDLTFTROPVEMVYFSAVSMPEPEFAAC 551
483 RIAYAKTSCLAVLDDLYDTHGSLDLDLKFSEAVRWDISVLDSVRDNLKVCFLGLYNT 542
552 RFLFTKVAQLQTVLDDMDYTYGTLDLDELKLFTEAVRWDLSFTENLPD-YMKLCYQIYYDI 610

```





193 LRGLPYDLPYIHLLOTKQERLAKLSREBIYAVPSPLLYSLEGIQDIVEWERIMEVQSQ 252  
QY  
183 DGSFLSPASTACVFWHTONKRCILHFLNVLKFGDYPCHYPLDLFERLWAVDVTVERLG 242  
Db  
253 DGSFLSPASTACVFWHTGDAKCLDFLNSVMKFGNFVCLYPVDLLERLLIVDNIVRLG 312  
QY  
243 IDRYFKKEIKESLDYVRYWDAERGVCWACNPIPDVDDTAMGLRIILRHGYNVSSDVLE 302  
Db  
313 IYRHFKEIKEALDYVYRHN-ERGIWGRLNPIADLETTALGFLRLRLHRYNVSPAIFD 371  
QY  
303 NFRDEKGFPCFAGOTQIGTDMNLNLYRCQVCPGCEKIMEEAKTFTTNHLQNALAKNNA 362  
Db  
372 NFKDANGKFCSTGQFNKDVASMLNLYRASQLAPFGENILDEAKSPATKYLRLEKSET 431  
QY  
363 FDKWAVKDLPGVEYAIKYPWHRSMPLREARSYIEQFGSNDVWLGKTVYKMLYVSNEKY 422  
Db  
432 SSANNKQNLQSEIKYALKTSWASVPRVEAKRYCQVYRDPYARIACVYKLPYVNNKEF 491  
QY  
423 LELAKLDFNMVQALHQKETCHIVSWRESGNDLTFTQRQPVEMVFSVAVSMPEPEFAAC 482  
Db  
492 LELKLDNFNIIQSTHQSEMKNVTSWFRDSDGLPFTFARERPLEFYFLVAAGTYEPQYAKC 551  
QY  
483 RIAYAKTSLAVIILDDLYDTHGSLDDKLFGSEAVRRWDISVLDSVRDNLKVCFLGLYNT 542  
Db  
552 RFLFTKVACLQTVLDDMYDTYGTLDLKLFTAVRRWDLSTENLPD-YMKLCYIYDI 610  
QY  
543 VNGFGKDLQKQGRDVLGRLKRWEGLLASYTKEAWSAAKYVPTFNBYENAKVSIALA 602  
Db  
611 VHEVANEAEKQGRELVSFRRKGMEDYLLGYVEEAELAAEYVPTLDEYIKNGITSIGOR 670  
QY  
603 TVVLNSI-FFTGELLPDYILOQVDL--RSKFLHLVSLTGRLLINDTKTYOABERNRGEVSS 659  
Db  
671 ILLLSGLVIMDQQLLSQALEKVDYPGRVLTSLNSLSRLADDTKYTAKKARGELASS 730  
QY  
660 VOCYMRNPECTERREALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNFTARVMOLF 719  
Db  
731 IECYMKDHPCTEERREALDHIYSILEPAVKELTREFLKP-DDVPPACKQMLFEETRTVTWVI 789  
QY  
720 YMYRDGFGISDKEMKHVSRTLFDPV 745  
Db  
790 PKDGDGFGVSKLEVKHKECLIEPL 815

Search completed: August 17, 2004, 21:03:50  
Job time : 18.1272 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 21:02:37 ; Search time 40.5485 Seconds  
(without alignments)  
5775.546 Million cell updates/sec

Title: US-10-041-007-39  
Perfect score: 3972  
Sequence: 1 MFQSMGDTNPSAYDTAW.....GISDKEMKHVSRTLPDPA 746

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3972	100.0	746	13	US-10-041-007-39
2	3972	100.0	795	13	US-10-041-007-37
3	3972	100.0	814	13	US-10-041-007-35
4	3972	100.0	873	12	US-10-041-018-398
5	3972	100.0	873	13	US-10-041-007-2
6	3972	100.0	873	13	US-10-041-007-33
7	2735.5	68.9	868	9	US-09-887-586A-56
8	2735.5	68.9	868	9	US-09-903-012-56
9	2735.5	68.9	868	12	US-09-900-797-56
10	2735.5	68.9	868	12	US-09-893-820-56
11	2735.5	68.9	868	12	US-10-041-018-388
12	2735.5	68.9	868	13	US-10-041-007-4
13	2050.5	51.6	782	9	US-09-887-586A-46
14	2050.5	51.6	782	9	US-09-903-012-46
15	2050.5	51.6	782	10	US-09-900-797-46

16	2050.5	51.6	782	12	US-09-893-820-46	Sequence 46, Appl
17	2050.5	51.6	782	13	US-10-041-007-16	Sequence 16, Appl
18	2050.5	51.6	782	14	US-10-025-145A-16	Sequence 16, Appl
19	2050.5	51.6	817	13	US-10-041-007-15	Sequence 15, Appl
20	1901	47.9	862	9	US-09-887-586A-44	Sequence 44, Appl
21	1901	47.9	862	9	US-09-903-012-44	Sequence 44, Appl
22	1901	47.9	862	10	US-09-900-797-44	Sequence 44, Appl
23	1901	47.9	862	12	US-09-893-820-44	Sequence 44, Appl
24	1901	47.9	862	12	US-10-041-018-386	Sequence 386, App
25	1901	47.9	862	13	US-10-041-007-41	Sequence 41, Appl
26	1382	34.8	787	12	US-10-041-018-385	Sequence 385, App
27	1363.5	34.3	815	9	US-09-371-307-88	Sequence 88, Appl
28	1363.5	34.3	815	15	US-10-401-321-88	Sequence 88, Appl
29	1293.5	32.6	823	12	US-10-041-018-396	Sequence 396, App
30	1281	32.3	840	16	US-10-437-963-182585	Sequence 182585,
31	1275.5	32.1	801	12	US-10-041-018-394	Sequence 394, App
32	1252	31.5	581	9	US-09-887-586A-48	Sequence 48, Appl
33	1252	31.5	581	9	US-09-903-012-48	Sequence 48, Appl
34	1252	31.5	581	10	US-09-900-797-48	Sequence 48, Appl
35	1252	31.5	581	12	US-09-893-820-48	Sequence 48, Appl
36	1252	31.5	581	13	US-10-041-007-18	Sequence 18, Appl
37	1215.5	30.6	630	14	US-10-025-145A-67	Sequence 67, Appl
38	1206	30.4	577	14	US-10-025-145A-18	Sequence 18, Appl
39	1194.5	30.1	637	9	US-09-887-586A-58	Sequence 58, Appl
40	1194.5	30.1	637	9	US-09-903-012-58	Sequence 58, Appl
41	1194.5	30.1	637	10	US-09-900-797-58	Sequence 58, Appl
42	1194.5	30.1	637	12	US-09-893-820-58	Sequence 58, Appl
43	1194.5	30.1	637	13	US-10-041-007-24	Sequence 24, Appl
44	1194.5	30.1	637	14	US-10-025-145A-6	Sequence 6, Appl
45	1193	30.0	637	13	US-10-041-007-28	Sequence 28, Appl

## ALIGNMENTS

### RESULT 1

US-10-041-007-39  
; Sequence 39, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P.T.  
; APPLICANT: Schepmann, Hala G.  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081US1  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US 60/259,881  
; PRIOR FILING DATE: 2001-01-05  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 39  
; LENGTH: 746  
; TYPE: PRT  
; ORGANISM: Ginkgo biloba  
US-10-041-007-39

Query Match 100.0%; Score 3972; DB 13; Length 746;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MFQSMGDTNPSAYDTAWARIPSIDGSGAPOTQTLQILNNQLPDGSGEGECIPLAY	60
Db	1	MFQSMGDTNPSAYDTAWARIPSIDGSGAPOTQTLQILNNQLPDGSGEGECIPLAY	60
Qy	61	DRVLNTLACLLTLTKIKWKGDIQVQKGVFVRKMEEMKDEADNHRPSGFEVFPFAMLDEA	120
Db	61	DRVLNTLACLLTLTKIKWKGDIQVQKGVFVRKMEEMKDEADNHRPSGFEVFPFAMLDEA	120
Qy	121	KSLGLDLPYHLPFISQIHQKQKQLQKIPNLVNLHQTALLYSLQGLQDVVDMQEIINLQ	180
Db	121	KSLGLDLPYHLPFISQIHQKQKQLQKIPNLVNLHQTALLYSLQGLQDVVDMQEIINLQ	180
Qy	181	SRDGSFLSSPASTACVPMHTQNKRLHFLNVLVLSKFGDYVPCHPYPLDLFRLNAVDIVR	240

Db	181	SRDGSFLSPASTACVFMHTQNKRCCLHFLNFVLSKFGDYVPCHPYPLDLFRLWAVDTVER	240
Qy	241	LGIDRYFKKEIKESLDVYVRYWDAERGVCWACNPIPDVDDTAMGLRILRLHGYNVSSDV	300
Db	241	LGIDRYFKKEIKESLDVYVRYWDAERGVCWACNPIPDVDDTAMGLRILRLHGYNVSSDV	300
Qy	301	LENFRDEKGGFFCFAGQTOIGVTDNLNLYRCSQVCFPEGEKIMEBAKFTTNHLQNALAKN	360
Db	301	LENFRDEKGGFFCFAGQTOIGVTDNLNLYRCSQVCFPEGEKIMEBAKFTTNHLQNALAKN	360
Qy	361	NAPDKWAVKKDLPGCEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWLKTYVKMLYYSNE	420
Db	361	NAPDKWAVKKDLPGCEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWLKTYVKMLYYSNE	420
Qy	421	KYLELAKLDNFMVQALHOKETQHIVSWRRSGGFNDLFTTRQRPVEMYFSVAVSMFEPEFA	480
Db	421	KYLELAKLDNFMVQALHOKETQHIVSWRRSGGFNDLFTTRQRPVEMYFSVAVSMFEPEFA	480
Qy	481	ACRIAYAKTSCLAVILDDLYDTHGSLDDDLKLFSEAVRRWDISVLDSVRDNLKVCFLGLY	540
Db	481	ACRIAYAKTSCLAVILDDLYDTHGSLDDDLKLFSEAVRRWDISVLDSVRDNLKVCFLGLY	540
Qy	541	NTVNGFGKGLKEQGRDVLGYLRKVWEGLLASYTEKEAWSAAKVPTFNEVENAKVSIA	600
Db	541	NTVNGFGKGLKEQGRDVLGYLRKVWEGLLASYTEKEAWSAAKVPTFNEVENAKVSIA	600
Qy	601	LATVVLNSIFFTGBELLPDYILQQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNRGELVSSV	660
Db	601	LATVVLNSIFFTGBELLPDYILQQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNRGELVSSV	660
Qy	661	QCYMRENPECTEEALSHPVYGIIDNALKELWELANPASNAPLCVRRLLFNTARVMQLFY	720
Db	661	QCYMRENPECTEEALSHPVYGIIDNALKELWELANPASNAPLCVRRLLFNTARVMQLFY	720
Qy	721	MYRDGFGISDKEMKDHVSRTILFDPVA	746
Db	721	MYRDGFGISDKEMKDHVSRTILFDPVA	746
RESULT 2			
US-10-041-007-37			
; Sequence 37, Application US/10041007			
; Publication No. US20020164736A1			
; GENERAL INFORMATION:			
; APPLICANT: Matsuda, Seiichi P.T.			
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase			
; FILE REFERENCE: P02081US1			
; CURRENT APPLICATION NUMBER: US/10/041,007			
; PRIOR FILING DATE: 2002-01-07			
; PRIOR APPLICATION NUMBER: US 60/259,881			
; NUMBER OF SEQ ID NOS: 41			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 37			
; LENGTH: 795			
; TYPE: PRT			
; ORGANISM: Ginkgo biloba			
US-10-041-007-37			
Query Match 100.0%; Score 3972; DB 13; Length 795;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MFQSMGDETNPSTAYTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGECEIFLAY	60
Db	50	MFQSMGDETNPSTAYTAWARIPSIDSGAPQFPQTLOWILNNQLPDGSGWGECEIFLAY	109
Qy	61	DRVLTACLLTLKINWKGDIQVKGVEFVRKHEEMKDEADNHRPSGFEVFPAMLDEA	120
Db	110	DRVLTACLLTLKINWKGDIQVKGVEFVRKHEEMKDEADNHRPSGFEVFPAMLDEA	169

```
QY 61 DRVLANTLACLLTLKIWNKGDIOVKGVFVRKHEEMKDEADNHRPSGFEVVPAMLDEA 120
DB 129 DRVLANTLACLLTLKIWNKGDIOVKGVFVRKHEEMKDEADNHRPSGFEVVPAMLDEA 188
QY 121 KSLGLDLPYHLPPISQIHQKROKQLQKIPLNVLNHNHTALLYSLEGLQDVDMQEIITNLQ 180
DB 189 KSLGLDLPYHLPPISQIHQKROKQLQKIPLNVLNHNHTALLYSLEGLQDVDMQEIITNLQ 248
QY 181 SRDGSPLSSPASTACVFMHTQNKRCCLHFLNVLNFKSGDYVPCHPYLDLPERLMAVDTVR 240
DB 249 SRDGSPLSSPASTACVFMHTQNKRCCLHFLNVLNFKSGDYVPCHPYLDLPERLMAVDTVR 308
QY 241 LGIDRYFKKEIKESLDYVYRYMDAERGVGWACNPIPDVDDTAMGLRIILRLHGYNVSSDV 300
DB 309 LGIDRYFKKEIKESLDYVYRYMDAERGVGWACNPIPDVDDTAMGLRIILRLHGYNVSSDV 368
QY 301 LENFRDEKDDFFCFAGQTOIGVTDNLNLRYCSQVCPGGEKIMEEAKTFTTNHLQNALAKN 360
DB 369 LENFRDEKDDFFCFAGQTOIGVTDNLNLRYCSQVCPGGEKIMEEAKTFTTNHLQNALAKN 428
QY 361 NAFDKWAVKVDLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWLGKTVYKMLYVSN 420
DB 429 NAFDKWAVKVDLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWLGKTVYKMLYVSN 488
QY 421 KYLELAKLDPNMQALHOKETQHIIVSWWRRESGNDLTFTRQRPVEMYFSVAVSMFPEPFA 480
DB 489 KYLELAKLDPNMQALHOKETQHIIVSWWRRESGNDLTFTRQRPVEMYFSVAVSMFPEPFA 548
QY 481 ACRIAYAKTSCLAVIDLDLYDTHGSLDDLLKLPSEAVRRWDISVLDSVRDNQKVCPLGLY 540
DB 549 ACRIAYAKTSCLAVIDLDLYDTHGSLDDLLKLPSEAVRRWDISVLDSVRDNQKVCPLGLY 608
QY 541 NTVNGFGDKGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSI 500
DB 609 NTVNGFGDKGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSI 568
QY 601 LATVVLNSIFFTGCELLPDYILOQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNRGEVSSV 560
DB 669 LATVVLNSIFFTGCELLPDYILOQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNRGEVSSV 728
QY 661 QCYMRENPECTEBEALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 720
DB 729 QCYMRENPECTEBEALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 788
QY 721 MYRDGFGISDKEMKOHVSRTLPDPA 746
DB 789 MYRDGFGISDKEMKOHVSRTLPDPA 814
```

## RESULT 4

```
US-10-041-018-398
; Sequence 398, Application US/10041018
; Publication No. US20040072323A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seichi P.T.
; APPLICANT: Hart, Elizabeth A.
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism
; FILE REFERENCE: P02080US1/10025547
; CURRENT APPLICATION NUMBER: US/10/041,018
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259880
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 413
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 398
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Ginkgo biloba
US-10-041-018-398
```

```
Query Match 100.0%; Score 3972; DB 12; Length 873;
Best Local Similarity 100.0%; Pred. No. 0;
```

```
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFQSMGDMGTNPSAYDTAWVARIPSIDGSGAPQFPOTLQWILNNQLPDGSGWGEBCIFLAY 60
DB 128 MFQSMGDMGTNPSAYDTAWVARIPSIDGSGAPQFPOTLQWILNNQLPDGSGWGEBCIFLAY 187
QY 61 DRVLANTLACLLTLKIWNKGDIOVKGVFVRKHEEMKDEADNHRPSGFEVVPAMLDEA 120
DB 188 DRVLANTLACLLTLKIWNKGDIOVKGVFVRKHEEMKDEADNHRPSGFEVVPAMLDEA 247
QY 121 KSLGLDLPYHLPPISQIHQKROKQLQKIPLNVLNHNHTALLYSLEGLQDVDMQEIITNLQ 180
DB 248 KSLGLDLPYHLPPISQIHQKROKQLQKIPLNVLNHNHTALLYSLEGLQDVDMQEIITNLQ 307
QY 181 SRDGSPLSSPASTACVFMHTQNKRCCLHFLNVLNFKSGDYVPCHPYLDLPERLMAVDTVR 240
DB 308 SRDGSPLSSPASTACVFMHTQNKRCCLHFLNVLNFKSGDYVPCHPYLDLPERLMAVDTVR 367
QY 241 LGIDRYFKKEIKESLDYVYRYMDAERGVGWACNPIPDVDDTAMGLRIILRLHGYNVSSDV 300
DB 368 LGIDRYFKKEIKESLDYVYRYMDAERGVGWACNPIPDVDDTAMGLRIILRLHGYNVSSDV 427
QY 301 LENFRDEKDDFFCFAGQTOIGVTDNLNLRYCSQVCPGGEKIMEEAKTFTTNHLQNALAKN 360
DB 428 LENFRDEKDDFFCFAGQTOIGVTDNLNLRYCSQVCPGGEKIMEEAKTFTTNHLQNALAKN 487
QY 361 NAFDKWAVKVDLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWLGKTVYKMLYVSN 420
DB 488 NAFDKWAVKVDLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWLGKTVYKMLYVSN 547
QY 421 KYLELAKLDPNMQALHOKETQHIIVSWWRRESGNDLTFTRQRPVEMYFSVAVSMFPEPFA 480
DB 548 KYLELAKLDPNMQALHOKETQHIIVSWWRRESGNDLTFTRQRPVEMYFSVAVSMFPEPFA 607
QY 481 ACRIAYAKTSCLAVIDLDLYDTHGSLDDLLKLPSEAVRRWDISVLDSVRDNQKVCPLGLY 540
DB 608 ACRIAYAKTSCLAVIDLDLYDTHGSLDDLLKLPSEAVRRWDISVLDSVRDNQKVCPLGLY 667
QY 541 NTVNGFGDKGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSI 600
DB 668 NTVNGFGDKGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSI 727
QY 601 LATVVLNSIFFTGCELLPDYILOQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNRGEVSSV 660
DB 728 LATVVLNSIFFTGCELLPDYILOQVDLRSKFLHLVSLTGRLLINDTKTYQAEARNRGEVSSV 787
QY 661 QCYMRENPECTEBEALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 720
DB 788 QCYMRENPECTEBEALSHVYGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 847
QY 721 MYRDGFGISDKEMKOHVSRTLPDPA 746
DB 848 MYRDGFGISDKEMKOHVSRTLPDPA 873
```

## RESULT 5

```
US-10-041-007-2
; Sequence 2, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seichi P.T.
; APPLICANT: Schepmann, Hala G.
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 873
; TYPE: PRT
```

```

; ORGANISM: Ginkgo biloba
US-10-041-007-2

Query Match      100.0%; Score 3972; DB 13; Length 873;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFQSMGDETNPSAYDTAWARIPSIDSGAPQPTQTLQWLNQLPDGSWGECIFLAY 60
Db 128 MFQSMGDETNPSAYDTAWARIPSIDSGAPQPTQTLQWLNQLPDGSWGECIFLAY 187

Qy 61 DRVLNTLACLLTLKWNKGDIOVQGVFVRKHEEMKDEADNHRPSGFEVFPAMLDEA 120
Db 188 DRVLNTLACLLTLKWNKGDIOVQGVFVRKHEEMKDEADNHRPSGFEVFPAMLDEA 247

Qy 121 KSLGLDLPYHLFFISQIHQKQKQLKQIPLNVLNHNHTALLYSLGLODQVVDWQEIITNLQ 180
Db 248 KSLGLDLPYHLFFISQIHQKQKQLKQIPLNVLNHNHTALLYSLGLODQVVDWQEIITNLQ 307

Qy 181 SRDGSFLSSPASTACVFMHTQNKCLHFLNVLKFGDYVPCHPDLDFERLWAVDTVER 240
Db 308 SRDGSFLSSPASTACVFMHTQNKCLHFLNVLKFGDYVPCHPDLDFERLWAVDTVER 367

Qy 241 LGIDRYFKKEIKESLDYVYRWDAERGVMARCNPIPDVDDTAMGLRLRLHGYNVSSDV 300
Db 368 LGIDRYFKKEIKESLDYVYRWDAERGVMARCNPIPDVDDTAMGLRLRLHGYNVSSDV 427

Qy 301 LENFRDEKDPFCFAGQTOIGVTNMLNLYRCSQVCFPGKEIMEEAKTFTTNHLQNALAKN 360
Db 428 LENFRDEKDPFCFAGQTOIGVTNMLNLYRCSQVCFPGKEIMEEAKTFTTNHLQNALAKN 487

Qy 361 NAFDKWAVKKDLPGVEVYAIKYPWHRSMRLEARSYIEQFGSNDVWLKTVYKMLYVNE 420
Db 488 NAFDKWAVKKDLPGVEVYAIKYPWHRSMRLEARSYIEQFGSNDVWLKTVYKMLYVNE 547

Qy 421 KYLELAKLDFNNVQALHOKETQHIYSWWRRESGFNDLFTTRQRPVEMYFSVAVSMFPEPA 480
Db 548 KYLELAKLDFNNVQALHOKETQHIYSWWRRESGFNDLFTTRQRPVEMYFSVAVSMFPEPA 607

Qy 481 ACRIAYAKTSCLAVIDLDLYDTHGSLDLDLKFSEAVRWMDISVLDSDVRDNQKVCFLGLY 540
Db 608 ACRIAYAKTSCLAVIDLDLYDTHGSLDLDLKFSEAVRWMDISVLDSDVRDNQKVCFLGLY 667

Qy 541 NTVNGFGDGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSIA 600
Db 728 LATVVLNSIFFTGELLPDYILQVVDLRSKFLHLVSLTGRRLINDTKTYQAERNRGLVSSV 787

Qy 601 LATVVLNSIFFTGELLPDYILQVVDLRSKFLHLVSLTGRRLINDTKTYQAERNRGLVSSV 660
Db 728 LATVVLNSIFFTGELLPDYILQVVDLRSKFLHLVSLTGRRLINDTKTYQAERNRGLVSSV 787

Qy 661 QCYMRENPECTEBEALSHVYIGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 720
Db 788 QCYMRENPECTEBEALSHVYIGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 847

Qy 721 MYRDGFGISDKEMKDHVSRTLDPDPA 746
Db 848 MYRDGFGISDKEMKDHVSRTLDPDPA 873

RESULT 6
US-10-041-007-33
; Sequence 33, Application US/10041007
; Publication No. US20020164736A1
; GENERAL INFORMATION:
; APPLICANT: Matsuda, Seiichi P.T.
; APPLICANT: Schepmann, Hala G
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase
; FILE REFERENCE: P02081US1
; CURRENT APPLICATION NUMBER: US/10/041,007
; CURRENT FILING DATE: 2002-01-07
; PRIOR APPLICATION NUMBER: US 60/259,881
; PRIOR FILING DATE: 2001-01-05

```

```

; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 873
; TYPE: PRT
; ORGANISM: Ginkgo biloba
US-10-041-007-33

Query Match      100.0%; Score 3972; DB 13; Length 873;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MFQSMGDETNPSAYDTAWARIPSIDSGAPQPTQTLQWLNQLPDGSWGECIFLAY 60
Db 128 MFQSMGDETNPSAYDTAWARIPSIDSGAPQPTQTLQWLNQLPDGSWGECIFLAY 187

Qy 61 DRVLNTLACLLTLKWNKGDIOVQGVFVRKHEEMKDEADNHRPSGFEVFPAMLDEA 120
Db 188 DRVLNTLACLLTLKWNKGDIOVQGVFVRKHEEMKDEADNHRPSGFEVFPAMLDEA 247

Qy 121 KSLGLDLPYHLFFISQIHQKQKQLKQIPLNVLNHNHTALLYSLGLODQVVDWQEIITNLQ 180
Db 248 KSLGLDLPYHLFFISQIHQKQKQLKQIPLNVLNHNHTALLYSLGLODQVVDWQEIITNLQ 307

Qy 181 SRDGSFLSSPASTACVFMHTQNKCLHFLNVLKFGDYVPCHPDLDFERLWAVDTVER 240
Db 308 SRDGSFLSSPASTACVFMHTQNKCLHFLNVLKFGDYVPCHPDLDFERLWAVDTVER 367

Qy 241 LGIDRYFKKEIKESLDYVYRWDAERGVMARCNPIPDVDDTAMGLRLRLHGYNVSSDV 300
Db 368 LGIDRYFKKEIKESLDYVYRWDAERGVMARCNPIPDVDDTAMGLRLRLHGYNVSSDV 427

Qy 301 LENFRDEKDPFCFAGQTOIGVTNMLNLYRCSQVCFPGKEIMEEAKTFTTNHLQNALAKN 360
Db 428 LENFRDEKDPFCFAGQTOIGVTNMLNLYRCSQVCFPGKEIMEEAKTFTTNHLQNALAKN 487

Qy 361 NAFDKWAVKKDLPGVEVYAIKYPWHRSMRLEARSYIEQFGSNDVWLKTVYKMLYVNE 420
Db 488 NAFDKWAVKKDLPGVEVYAIKYPWHRSMRLEARSYIEQFGSNDVWLKTVYKMLYVNE 547

Qy 421 KYLELAKLDFNNVQALHOKETQHIYSWWRRESGFNDLFTTRQRPVEMYFSVAVSMFPEPA 480
Db 548 KYLELAKLDFNNVQALHOKETQHIYSWWRRESGFNDLFTTRQRPVEMYFSVAVSMFPEPA 607

Qy 481 ACRIAYAKTSCLAVIDLDLYDTHGSLDLDLKFSEAVRWMDISVLDSDVRDNQKVCFLGLY 540
Db 608 ACRIAYAKTSCLAVIDLDLYDTHGSLDLDLKFSEAVRWMDISVLDSDVRDNQKVCFLGLY 667

Qy 541 NTVNGFGDGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSIA 600
Db 668 NTVNGFGDGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSIA 727

Qy 601 LATVVLNSIFFTGELLPDYILQVVDLRSKFLHLVSLTGRRLINDTKTYQAERNRGLVSSV 660
Db 728 LATVVLNSIFFTGELLPDYILQVVDLRSKFLHLVSLTGRRLINDTKTYQAERNRGLVSSV 787

Qy 661 QCYMRENPECTEBEALSHVYIGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 720
Db 788 QCYMRENPECTEBEALSHVYIGIIDNALKEINWELANPASNAPLCVRRLLFNTARVMQLFY 847

Qy 721 MYRDGFGISDKEMKDHVSRTLDPDPA 746
Db 848 MYRDGFGISDKEMKDHVSRTLDPDPA 873

RESULT 7
US-09-887-586A-56
; Sequence 56, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.

```

```

; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; PRIOR FILING DATE: 2001-06-22
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-887-586A-56

Query Match      68.9%; Score 2735.5; DB 9; Length 868;
Best Local Similarity 67.6%; Pred. No. 8.7e-257;
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;

QY 1 MFOSMGDGTNPSTAYTAWARIPSDGSCAPOTPTLOWILNNOLPDGSGWGECEIFLAY 50
DB 126 MFRCMGYGTNPSTAYTAWARIPAVDGSNDPHPEVWILQNLQKDGSGWGECEIFLAY 185
QY 61 DRVLNTLACLLTLKWNKGDIQVKGVEFVRKHEEMKDEADNHRPSGFEVFPAMLDEA 120
DB 186 DRILATLACIITLWRTGETVQKGIPTDVLVYALPTTLLYSLEGLQEIWDQKIMKLQ 245
QY 121 KSLGLDPLPHLPISQIHQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 180
DB 246 KILGLDPLDPLKQIIEKREAKKRIPTDVLVYALPTTLLYSLEGLQEIWDQKIMKLQ 305
QY 181 SRDGSFLSPASTACVPMHTONKRCCLHFLNPLVLSKFGDVVPCHPYDLPDLPERLWAVDIVER 240
DB 306 SKDGSFLSPASTAAVPMRTGNKKCLDPLNFKFNGHVPCHYDLPDLPERLWAVDIVER 365
QY 241 LGIDRYFKKEIKESLDVYRYMDAERGVMGWARCNPIPDVDDTAMGLRILRLHGYNVSSDV 300
DB 366 LGIDRHFKKEIKESLDVYRYMDAERGVMGWARCNPIPDVDDTAMGLRILRLHGYNVSSDV 424
QY 301 LENFRDEKDPFCFAGQOTQIGVTDNLNLYRCQVCPGKIMEAKTFTTNHQLNALAKN 360
DB 425 LKTRFDENGEBFCFLGQOTQIGVTDNLNLYRCQVCPGKIMEAKTFTTNHQLNALAKN 484
QY 361 NAFDKWAVKOLPGEVEYAIKYPWHRSMPLREARSIEQPGSNDVWLGKTVYKMLYSNE 420
DB 485 DAFDKWAFKKNIRGEVEYAIKYPWHRSMPLREARSIEQPGSNDVWLGKTVYKMLYSNE 544
QY 421 KYLELAKLDFNMVQALHOKETQHIWVSWRESGNDLFTTRQRPVEMYFVSVAVMFPEFA 480
DB 545 KYLELAKLDFNMVQALHOKETQHIWVSWRESGNDLFTTRQRPVEMYFVSVAVMFPEFA 604
QY 481 ACRIAYAKTSCLAVIIDLDDYTHGSLDLDLKPSEAVRRWDISVLDSVRDNLQKVCPLGLY 540
DB 605 KCREVYTKTSNFTVILDDYDAHGSLLDLKFLTESVKRWDLSLVDQM-PQOMKICFVGFY 663
QY 541 NTVNGFGKGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSTA 600

```

```

DB 842 MQGDGLTLSDHMEIKHEVKNCLFPQVA 868

RESULT 8
US-09-903-012-56
; Sequence 56, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starke, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012
; PRIOR FILING DATE: 2001-07-11
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-09-17
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-04-22
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 56
; LENGTH: 868
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-903-012-56

Query Match      68.9%; Score 2735.5; DB 9; Length 868;
Best Local Similarity 67.6%; Pred. No. 8.7e-257;
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;

QY 1 MFOSMGDGTNPSTAYTAWARIPSDGSCAPOTPTLOWILNNOLPDGSGWGECEIFLAY 60
DB 126 MFRCMGYGTNPSTAYTAWARIPAVDGSNDPHPEVWILQNLQKDGSGWGECEIFLAY 185
QY 61 DRVLNTLACLLTLKWNKGDIQVKGVEFVRKHEEMKDEADNHRPSGFEVFPAMLDEA 120
DB 186 DRILATLACIITLWRTGETVQKGIPTDVLVYALPTTLLYSLEGLQEIWDQKIMKLQ 245
QY 121 KSLGLDPLPHLPISQIHQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 180
DB 246 KILGLDPLDPLKQIIEKREAKKRIPTDVLVYALPTTLLYSLEGLQEIWDQKIMKLQ 305
QY 181 SRDGSFLSPASTACVPMHTONKRCCLHFLNPLVLSKFGDVVPCHPYDLPDLPERLWAVDIVER 240
DB 306 SKDGSFLSPASTAAVPMRTGNKKCLDPLNFKFNGHVPCHYDLPDLPERLWAVDIVER 365
QY 241 LGIDRYFKKEIKESLDVYRYMDAERGVMGWARCNPIPDVDDTAMGLRILRLHGYNVSSDV 300
DB 366 LGIDRHFKKEIKESLDVYRYMDAERGVMGWARCNPIPDVDDTAMGLRILRLHGYNVSSDV 424
QY 301 LENFRDEKDPFCFAGQOTQIGVTDNLNLYRCQVCPGKIMEAKTFTTNHQLNALAKN 360
DB 425 LKTRFDENGEBFCFLGQOTQIGVTDNLNLYRCQVCPGKIMEAKTFTTNHQLNALAKN 484
QY 361 NAFDKWAVKOLPGEVEYAIKYPWHRSMPLREARSIEQPGSNDVWLGKTVYKMLYSNE 420
DB 485 DAFDKWAFKKNIRGEVEYAIKYPWHRSMPLREARSIEQPGSNDVWLGKTVYKMLYSNE 544
QY 421 KYLELAKLDFNMVQALHOKETQHIWVSWRESGNDLFTTRQRPVEMYFVSVAVMFPEFA 480
DB 545 KYLELAKLDFNMVQALHOKETQHIWVSWRESGNDLFTTRQRPVEMYFVSVAVMFPEFA 604
QY 481 ACRIAYAKTSCLAVIIDLDDYTHGSLDLDLKPSEAVRRWDISVLDSVRDNLQKVCPLGLY 540
DB 605 KCREVYTKTSNFTVILDDYDAHGSLLDLKFLTESVKRWDLSLVDQM-PQOMKICFVGFY 663
QY 541 NTVNGFGKGLKEQGRDVLGYLRKWEGLLASVYTKAEWSAAKYVPTFNEYVENAKVSTA 600

```

Db	664	NTFNDIAKEGRERQGRDVLGYIQNVKQLEAYTKEAEWSEAKVPSFNEIENASVIA	723
Qy	601	LATVVLNSIFFTGELLPDYILQOVDLRSKFLHLVSLTGRINDTKTYQAERNRGELVSV	660
Db	724	LGTVVLISALFTGEVLDEVLSKIDRESRFQLMGLTGRLVNDTKTYQAERGQGEVASAI	783
Qy	661	QCYMRENPECTEERALSHTVYGIIDNALKELWELANPASNAPLCVRLLFNTARVMOLFY	720
Db	784	QCYMKDHPKISEEALQHVYSVMENALEELNREFFV--NKIPDIYKRLVFTARIMOLFY	841
Qy	721	MYRDFGIS-DKEMKDVSRFLDPVA	746
Db	842	MQDGLTSLSHDMEIKHVKNCLFQFVA	868
RESULT 9			
US-09-900-797-56			
; Sequence 56, Application US/09900797			
; Publication No. US20030087406A1			
; GENERAL INFORMATION:			
; APPLICANT: Chappell, Joseph			
; APPLICANT: No. US20030087406A11, Joseph P.			
; APPLICANT: Starks, Courtney M.			
; APPLICANT: Manna, Kathleen R.			
; TITLE OF INVENTION: SYNTHASES			
; FILE REFERENCE: 07678-025001			
; CURRENT APPLICATION NUMBER: US/09/900,797			
; PRIOR FILING DATE: 2001-07-06			
; PRIOR APPLICATION NUMBER: US/09/398,395			
; PRIOR FILING DATE: 1999-09-17			
; PRIOR APPLICATION NUMBER: 60/130,628			
; PRIOR FILING DATE: 1999-04-22			
; PRIOR APPLICATION NUMBER: 60/150,262			
; PRIOR FILING DATE: 1999-08-23			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 56			
; LENGTH: 868			
; TYPE: PRT			
; ORGANISM: Abies grandis			
US-09-900-797-56			
Query Match 68.9%; Score 2735.5; DB 10; Length 868;			
Best Local Similarity 67.6%; Pred. No. 8.7e-257;			
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;			
Qy	1	MFQSMGDETNPSAYDTAWVARIPSDGSGAPQFPQTQLQWILNNQLPDGSGWGEICIFLAY	60
Db	126	MFRCMGYGETNFSAYDTAWVARIPAVDGSNDPHFPETVEWILQNLKDGSGWGEFYFLAY	185
Qy	61	DRVLTNLTACLLTKIWNKGDIOVKQGEVFRKHEEMKDEADNHRPSGEFVFPAMLDEA	120
Db	186	DRILATLACIIITLTWRTGETQVQKIEFFRTQAGKMEDEADSHRPSGEFVFPAMLKEA	245
Qy	121	KSIGLDLPVLPISQIHOQKROKLOKIPLVNLHHTALLYSLEGLOVDVWQEIETNLQ	180
Db	246	KILGLDLPVLPFLKQIIEKREAKLRIPDVLVYALPTTLLYSLEGLOEIVDQKIMKLQ	305
Qy	181	SRDGSFLSSPASTACVFMHTQNKRCFLHFLNVLKSGFDVYVPCHPYLDLFLERLWAVDTVER	240
Db	306	SKDGSFLSSPASTAAVFMRTGNKKCLDFNLVFLKFGNHPVCHYPYLDLFLERLWAVDTVER	365
Qy	241	LIGDRYFKKEIKESLDYVRYMDAERGCVGWARCNIPDVDDTAMGLRIILRHGYNVSSDV	300
Db	366	LGHDRHFKEEIKALDYVYVSHWD- ERGIGWARENPVDDIDDTANGRLRIILRHGYNVSSDV	424
Qy	301	LENFRDEKGFCCFAGQTOIGVTDNLNLVRCQVCFPEKIMEEAKTFTTNHLONALAKN	360
Db	425	LKTFRDENGEFCFLGQTQGVTDMLNVNRCSHVSFPGETIMEEAKLCTERYLRNALENV	484
Qy	361	NAPDKWAVKDLGEVEYALKYPWHRSMPLREARSYIETQGSNDVWLGKTYVMMPYISNE	420

Db	485	DAFDKWAFFKKNIRGEVEYALKYPWHKSMPLREARSYIETQGPDDVWLGKTYVMMPYISNE	544
Qy	421	KYLELAKLDENMQALHOKETOHLVSWWRSGFNDLFTTRQRPVEMYFSFVAVSMFEPEFA	480
Db	545	KYLELAKLDENKQVSIHQTELQDLRRWKSGGFDLNFTRERVTEIYFSPASPIFEPEFS	604
Qy	481	ACRIAYAKTSCLAVILDLDYTHGSLDDLKLFSAVNRWDLSVLDNRDNLKVCFLGLY	540
Db	605	KCREVYTKTSNFTVILDDLYDAHGSLLDLKLFTESVKRWDLSLVDQM-PQMKICFVGFY	663
Qy	541	NTVNGFOGKGLKEQGRDVLGYLKRKWEGLLASIYTKAEWSAAKYVPTFNEVVENAKVSIA	600
Db	664	NTFNDIAKEGRERQGRDVLGYIQNVKQLEAYTKEAEWSEAKVPSFNEIENASVIA	723
Qy	601	LATVVLNSIFFTGELLPDYILQOVDLRSKFLHLVSLTGRINDTKTYQAERNRGELVSV	660
Db	724	LGTVVLISALFTGEVLDEVLSKIDRESRFQLMGLTGRLVNDTKTYQAERGQGEVASAI	783
Qy	661	QCYMRENPECTEERALSHTVYGIIDNALKELWELANPASNAPLCVRLLFNTARVMOLFY	720
Db	784	QCYMKDHPKISEEALQHVYSVMENALEELNREFFV--NKIPDIYKRLVFTARIMOLFY	841
Qy	721	MYRDFGIS-DKEMKDVSRFLDPVA	746
Db	842	MQDGLTSLSHDMEIKHVKNCLFQFVA	868
RESULT 10			
US-09-893-820-56			
; Sequence 56, Application US/09893820			
; Publication No. US20040053386A1			
; GENERAL INFORMATION:			
; APPLICANT: Chappell, Joseph			
; APPLICANT: No. US20040053386A11, Joseph P.			
; APPLICANT: Starks, Courtney M.			
; APPLICANT: Manna, Kathleen R.			
; TITLE OF INVENTION: SYNTHASES			
; FILE REFERENCE: 07678-025001			
; CURRENT APPLICATION NUMBER: US/09/893,820			
; CURRENT FILING DATE: 2001-06-28			
; PRIOR APPLICATION NUMBER: US/09/398,395A			
; PRIOR FILING DATE: 1999-09-17			
; PRIOR APPLICATION NUMBER: 60/100,993			
; PRIOR FILING DATE: 1998-09-18			
; PRIOR APPLICATION NUMBER: 60/130,628			
; PRIOR FILING DATE: 1999-04-22			
; PRIOR APPLICATION NUMBER: 60/150,262			
; PRIOR FILING DATE: 1999-08-23			
; NUMBER OF SEQ ID NOS: 58			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 56			
; LENGTH: 868			
; TYPE: PRT			
; ORGANISM: Abies grandis			
US-09-893-820-56			
Query Match 68.9%; Score 2735.5; DB 12; Length 868;			
Best Local Similarity 67.6%; Pred. No. 8.7e-257;			
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;			
Qy	1	MFQSMGDETNPSAYDTAWVARIPSDGSGAPQFPQTQLQWILNNQLPDGSGWGEICIFLAY	60
Db	126	MFRCMGYGETNFSAYDTAWVARIPAVDGSNDPHFPETVEWILQNLKDGSGWGEFYFLAY	185
Qy	61	DRVLTNLTACLLTKIWNKGDIOVKQGEVFRKHEEMKDEADNHRPSGEFVFPAMLDEA	120
Db	186	DRILATLACIIITLTWRTGETQVQKIEFFRTQAGKMEDEADSHRPSGEFVFPAMLKEA	245
Qy	121	KSIGLDLPVLPISQIHOQKROKLOKIPLVNLHHTALLYSLEGLOVDVWQEIETNLQ	180
Db	246	KILGLDLPVLPFLKQIIEKREAKLRIPDVLVYALPTTLLYSLEGLOEIVDQKIMKLQ	305
Qy	181	SRDGSFLSSPASTACVFMHTQNKRCFLHFLNVLKSGFDVYVPCHPYLDLFLERLWAVDTVER	240



Db 306 SKGSGFLSPASTAAVFMRTGNKCLDLFLNFKFNGHVPCHPYLDLPERLWADVTVR 365  
Qy 241 LGIDRYFKKIBKESLDYVYRYWDAERGVCWRCNPIDVDDTAMGLRIILRLHGYNVSSDV 300  
Db 366 LGIDRHFKEEIKALDYVYSHWD-ERGIWARENPVDDDDTAMGLRIILRLHGYNVSSDV 424  
Qy 301 LENFRDEKDFPCFAGOTQIGVTDNLNLYRCQVCPGKEKIMEEAKTFTTNHLQNALAKN 360  
Db 425 LKTFRDEGEPFCFLGTQGTQGTVDMLNVNRCVSHVSPGGETIMEEAKLCTERYLNALENV 484  
Qy 361 NAFDKWAVKDLPGGEVYAIKYPWHRSMPLRLARSYIEOFGSNDVWLGKTVYKMLVYSNE 420  
Db 485 DAFDKWAFKKNRGEVEYALKYPWHSMPRLARSYIENYGGDDVWLGKTVYKMLVYSNE 544  
Qy 421 KYLELAKLDFNMVQALHOKETQHIYVSWRESGNDLFTTRQRPVEMYFVSVMFPEPEPA 480  
Db 545 KYLELAKLDFNMVQALHOKETQHIYVSWRESGNDLFTTRQRPVEMYFVSVMFPEPEPE 504  
Qy 481 ACRIAYAKTSCLAVIDLDLYTHGSLDLDLKLSEAVRWDDISVLSVRNQLKVCPLGLY 540  
Db 605 KCREVYTKTSNFTVLDLDDYDAHGSLLDLDLKLFTESVWRWDLSDVDM-QPOMKICPVGFY 563  
Qy 541 NTVNGFGDKLKEQGRDVLGYLRKWKWEGLLASVYTKAEAWSAKYVPTENEYVENAKVSTA 500  
Db 664 NTFNDIAKEGRGRQGRDVLGYIQNVKVOLEAYTKAEAWSEAKYVPSFNEYIENASVSTA 723  
Qy 601 LATVVLNSIFFTGELLPYDYLQOVDLRSKFLHLVSLTGRINDTKTYQAEARNRGLVSSV 560  
Db 724 LGTVVLISALFTGEVLTDVLSKIDRESRFQLMGLTGLVNDTKTYQAEARGQGEVASAI 783  
Qy 661 QCYMRNPECTBEERASHVYIGIIDNALKEINWELANPASNAFLCVRRLLFNTARVMQLPY 720  
Db 784 QCYMRDHPKISEEALQHVYSVMENALBELNREFVN--NKIPDIYKRLVFETARIMQLPY 341  
Qy 721 MYRDGFGIS-DKEMKHVSRTLPDPA 746  
Db 842 MOGDGLTSHDMEIKHVKNCFLQFQVPA 868

RESULT 11  
US-10-041-018-388  
; Sequence 388, Application US/10041018  
; Publication No. US20040072323A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P. T.  
; TITLE OF INVENTION: Diterpene-Producing Unicellular Organism  
; FILE REFERENCE: P02080USI/10025547  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2002-01-07  
; NUMBER OF SEQ ID NOS: 413  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 388  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Abies grandis  
US-10-041-018-388  
Query Match 68.9%; Score 2735.5; DB 12; Length 868;  
Best Local Similarity 67.6%; Pred. No. 8.7e-257;  
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;

Qy 1 MFQSGDGETNPSAYDTAWARIPSIDGSGAQFPQTLOWILNNQLPDGSGWCECFLAY 50  
Db 126 MFRCMGYGETNPSAYDTAWARIPAVDGSNDNPHFPETVETWILQNLKDGSGWCEGFYFLAY 185  
Qy 61 DRVLNTLACLLTKTWKNGDIOVQGVRFVRKHMEDKDEADNHRPSGFEVVPFPAWLDEA 120  
Db 186 DRILATLACIITLTWRTGETQVQGIETFTQAGMEDEADSHRPSGFEIIVPFAWLKEA 245

Qy 121 KSLGLDLPYHLPIFSQIHOKROKKLOKIPLVNLVHNHOTALLYSLEGLQDVVDHQBETNLQ 180  
Db 246 KTLGLDLPYDLFPKQIIEKREAKLRITDVLALPTLLYSLEGLQETVDWQKMKLQ 305  
Qy 181 SRDGSFLSPASTACVFMHTONKRCILHFLNPLVLSKFGDVVPCHPYLDLPERLWADVTVR 240  
Db 306 SKGSGFLSPASTAAVFMRTGNKCLDLFLNFKFNGHVPCHPYLDLPERLWADVTVR 365  
Qy 241 LGIDRYFKKIBKESLDYVYRYWDAERGVCWRCNPIDVDDTAMGLRIILRLHGYNVSSDV 300  
Db 366 LGIDRHFKEEIKALDYVYSHWD-ERGIWARENPVDDDDTAMGLRIILRLHGYNVSSDV 424  
Qy 301 LENFRDEKDFPCFAGOTQIGVTDNLNLYRCQVCPGKEKIMEEAKTFTTNHLQNALAKN 360  
Db 425 LKTFRDEGEPFCFLGTQGTQGTVDMLNVNRCVSHVSPGGETIMEEAKLCTERYLNALENV 484  
Qy 361 NAFDKWAVKDLPGGEVYAIKYPWHRSMPLRLARSYIEOFGSNDVWLGKTVYKMLVYSNE 420  
Db 485 DAFDKWAFKKNRGEVEYALKYPWHSMPRLARSYIENYGGDDVWLGKTVYKMLVYSNE 544  
Qy 421 KYLELAKLDFNMVQALHOKETQHIYVSWRESGNDLFTTRQRPVEMYFVSVMFPEPEPA 480  
Db 545 KYLELAKLDFNMVQALHOKETQHIYVSWRESGNDLFTTRQRPVEMYFVSVMFPEPEPE 604  
Qy 481 ACRIAYAKTSCLAVIDLDLYTHGSLDLDLKLSEAVRWDDISVLSVRNQLKVCPLGLY 540  
Db 605 KCREVYTKTSNFTVLDLDDYDAHGSLLDLDLKLFTESVWRWDLSDVDM-QPOMKICPVGFY 663  
Qy 541 NTVNGFGDKLKEQGRDVLGYLRKWKWEGLLASVYTKAEAWSAKYVPTENEYVENAKVSTA 600  
Db 664 NTFNDIAKEGRGRQGRDVLGYIQNVKVOLEAYTKAEAWSEAKYVPSFNEYIENASVSTA 723  
Qy 601 LATVVLNSIFFTGELLPYDYLQOVDLRSKFLHLVSLTGRINDTKTYQAEARNRGLVSSV 660  
Db 724 LGTVVLISALFTGEVLTDVLSKIDRESRFQLMGLTGLVNDTKTYQAEARGQGEVASAI 783  
Qy 661 QCYMRNPECTBEERASHVYIGIIDNALKEINWELANPASNAFLCVRRLLFNTARVMQLPY 720  
Db 784 QCYMRDHPKISEEALQHVYSVMENALBELNREFVN--NKIPDIYKRLVFETARIMQLPY 841  
Qy 721 MYRDGFGIS-DKEMKHVSRTLPDPA 746  
Db 842 MOGDGLTSHDMEIKHVKNCFLQFQVPA 868

RESULT 12  
US-10-041-007-4  
; Sequence 4, Application US/10041007  
; Publication No. US20020164736A1  
; GENERAL INFORMATION:  
; APPLICANT: Matsuda, Seiichi P. T.  
; TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
; FILE REFERENCE: P02081USI  
; CURRENT APPLICATION NUMBER: US/10/041,007  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR FILING DATE: 2002-01-07  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 4  
; LENGTH: 868  
; TYPE: PRT  
; ORGANISM: Abies grandis  
US-10-041-007-4  
Query Match 68.9%; Score 2735.5; DB 13; Length 868;  
Best Local Similarity 67.6%; Pred. No. 8.7e-257;  
Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;

Qy 1 MFQSGDGETNPSAYDTAWARIPSIDGSGAQFPQTLOWILNNQLPDGSGWCECFLAY 60  
Db 126 MFRCMGYGETNPSAYDTAWARIPAVDGSNDNPHFPETVETWILQNLKDGSGWCEGFYFLAY 185

```

QY 61 DRVLNTLACLLTTLKINWKGDIQVOKGVFVRKHEEMKMDADNHRPSGFEVVRPAMLEA 120
DB 186 DRILATLACITITLMTGTGTQVOKGLEFFRTQKWEDEADSHRPSGFEIVFPAMLEA 245
QY 121 KSLGLDLPYHLPFTSIHQKROKQKIKIPNLVNLHHTALLYSLEGQDVVDWQEIITNLQ 180
DB 246 KILGLDLPYDLPLFKQIIEKEAKLRIPDTDLVALPTLLYSLEGQEIWDQKIMKIQ 305
QY 181 SRGSLFSSPASTACVFMHTQKCLHPLNVLKFGDYVPCHPYLDLPERLWADVTVER 240
DB 306 SKDGSFLSSPASTAAVFMRTGNKCLDFLNVLKFGNHNVPCHPYLDLPERLWADVTVER 365
QY 241 LGIDRYFKRTIKESLDYVYRWDAERGVMGWARCNPIPDVDDTAMGLRLILRHGYNVSSDV 300
DB 366 LGIDRHFKKEEIKALDYVYSHWD-ERGIGWARENVPDIDDTAMGLRLILRHGYNVSSDV 424
QY 301 LENFRDEKDPFCFAGOTQIGVTDLNLRYCSQVCFGEKIMEBAKFTTNHLQNALAKN 360
DB 425 LKTFRDENGEEFCFLGQTQRTGVTDLNLVNRCSHVSFFGETIMERAUKCTERYLNALENV 484
QY 361 NAFDKWAKDLPOEVEYAIKYPHRSMPRLARSYIIOFGSNDVWLGKTYVKMLYSNE 420
DB 485 DAFDKWAFKQNIERGEVEYALKYPHKSMPRLARSYIENYGFDDVWLGKTYVMMPYSNE 544
QY 421 KYLELAKLDFNMVQALHOKETQHIIVSWWRESGENDLTFTROPVEMYSFVAVSMFEPEFA 480
DB 545 KYLELAKLDFNMVQALHOKETQHIIVSWWRESGENDLTFTROPVEMYSFVAVSMFEPEFA 604
QY 481 ACRIAYAKTSCLAVIDLDYTHGSLDDLKLFSEAVRWDISVLDSVRDNLQKVCFLGLY 540
DB 605 KCREVYTKTSNFTVLLDLYDAHGSDDLKLFSEAVRWDISVLDSVRDNLQKVCFLGLY 663
QY 541 NTVNGFGDGLKEQGRDVLGYLRKWEGLLASYTEKEAWSAAKYVPTNEVEYENAKVSIA 600
DB 664 NTFNDIAKEGRERQGRDVLGYLRKWEGLLASYTEKEAWSAAKYVPTNEVEYENAKVSIA 723
QY 601 LATVVLNSIPTGELLPPYILQVVDLRSKFLHVLSTGLRLINDTKTYQAEARNRGLVSV 660
DB 724 LGTVVLNSIPTGELLPPYILQVVDLRSKFLHVLSTGLRLINDTKTYQAEARNRGLVSV 783
QY 661 QCYMRENPECTEEALSHVYGIIDNALKELNWLANPASNAPLCVRRLLFNTARVMQLFY 720
DB 784 QCYMKDHPKISEEALQHVYSWMENALBELNREFVN--NKIPDIYKRLVFTETARIMQLFY 841
QY 721 MYRDGFGIS-DKEMKDHVSRTILFDPVA 746
DB 842 MQDGLTLSDHMEIKHEVKNCLFQFVA 868

```

```

RESULT 13
US-09-887-586A-46
; Sequence 46, Application US/09887586A
; Patent No. US20020094556A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094556A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/887,586A
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 782

```

```

; TYPE: PRT
; ORGANISM: Abies grandis
US-09-887-586A-46
Query Match 51.6%; Score 2050.5; DB 9; Length 782;
Best Local Similarity 50.8%; Pred. No. 4.4e-190; Indels 9; Gaps 7;
Matches 379; Conservative 152; Mismatches 206;
QY 6 GDGSET--NPSAYDTAWVARIPSIDSGAPQFPQTLOWILNNQLPDGSGEBCIFLAYDRV 63
DB 38 GDGESMITPSAYDTAWVARIPSIDSGAPQFPQTLOWILNNQLPDGSGEBCIFLAYDRV 97
QY 64 LNTIACLLTTLKINWKGDIQVOKGVFVRKHEEMKMDADNHRPSGFEVVRPAMLEA 122
DB 98 LATLSCLVLLKMNVDLQVQEGIEFIKSNELVKDETDQDSLVDFPIIPSLREQAQS 157
QY 123 LGIDLPHYLPFTSIHQKROKQKIKIPNLVNLHHTALLYSLEGQDVVDWQEIITNLQSR 182
DB 158 LRLGLPYDLPLHYLLOTTKQERLAKLSREEIYAVSPPLYSLEGQDIWENRIMEVQSQ 217
QY 183 DGSFLSSPASTACVFMHTQKCLHPLNVLKFGDYVPCHPYLDLPERLWADVTVERLG 242
DB 218 DGSFLSSPASTACVFMHTQKCLHPLNVLKFGDYVPCHPYLDLPERLWADVTVERLG 277
QY 243 IDRYFKRTIKESLDYVYRWDAERGVMGWARCNPIPDVDDTAMGLRLILRHGYNVSSDVLE 302
DB 278 IYRHFPEKEIKALDYVYSHWD-ERGIGWARENVPDIDDTAMGLRLILRHGYNVSSDVLE 336
QY 303 NFRDEKDPFCFAGOTQIGVTDLNLRYCSQVCFGEKIMEBAKFTTNHLQNALAKNNA 362
DB 337 NFKDANGKFI CSTGQFNKDVASMLNLYRASQAFEGENILDEAKSFATKYLEALEKSET 396
QY 363 FDKWAKVKKDLPOEVEYAIKYPHRSMPRLARSYIIOFGSNDVWLGKTYVKMLYSNEKY 422
DB 397 SSANNKQNLQSEIKYALKTSWASVPRVEAKRYCQVYRPDYARIKACVYKLPYVNNKEF 456
QY 423 LELAKLDFNMVQALHOKETQHIIVSWWRESGENDLTFTROPVEMYSFVAVSMFEPEFAAC 482
DB 457 LELGKLDENIIQSIHQEEMKNVTSWFRDGLPLFTFAERPLFEYFLVAAGTYEPOYAKC 516
QY 483 RIAYAKTSCLAVIDLDYTHGSLDDLKLFSEAVRWDISVLDSVRDNLQKVCFLGLYNT 542
DB 517 RPLFTKVAACLTQVLDMDYDTGTDLDELKLTFAVVRWDLSTFENLPD-YMKLCYQIYDI 575
QY 543 VNGFGDGLKEQGRDVLGYLRKWEGLLASYTEKEAWSAAKYVPTNEVEYENAKVSIALA 602
DB 576 VHEVAWEAEKEQRELVSFFRKGWEDYLLGYVEEAELAAEYVPTLDYIKNGITSIGOR 635
QY 603 TVVLNSI-PFTGELLPPYILQVVDL--RSKFLHVLSTGLRLINDTKTYQAEARNRGLVSS 659
DB 636 ILLSGVLIMDQGLLSQEALEKVDYPRGRVLTSLNLSIRLADDTTKTYKAEKARGELASS 695
QY 660 QCYMRENPECTEEALSHVYGIIDNALKELNWLANPASNAPLCVRRLLFNTARVMQLF 719
DB 696 IECYMKDHPKISEEALQHVYSWMENALBELNREFVN--DDVPFFACKMFLFEETRTMTVI 754
QY 720 YMYRDGFGIS-DKEMKDHVSRTILFDPV 745
DB 755 FKDGDFGVSKLEVKHKEICLIEPL 780

```

```

RESULT 14
US-09-903-012-46
; Sequence 46, Application US/09903012
; Patent No. US20020094557A1
; GENERAL INFORMATION:
; APPLICANT: Chappell, Joseph
; APPLICANT: No. US20020094557A1, Joseph P.
; APPLICANT: Starks, Courtney M.
; APPLICANT: Manna, Kathleen R.
; TITLE OF INVENTION: SYNTHASES
; FILE REFERENCE: 07678-025001
; CURRENT APPLICATION NUMBER: US/09/903,012

```

```
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-903-012-46

Query Match      51.6%; Score 2050.5; DB 9; Length 782;
Best Local Similarity 50.8%; Pred. No. 4.4e-190;
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;

QY 6 GDGET--NPSAYDTAWARIPISIDSGAPQFPQTQWILNNQLPDGSGWGEICIFLAYDRV 53
Db 38 GDGESMITPSAYDTAWARVPAIDGSARPQFPQTQWILNNQLKDGSGWGIQSHFLSDDL 37
QY 64 LNTLACLLTLKWNKGDIOVQKGVFVRKHEMKDEAD-NHRPSGFVFPVPMLEAKS 122
Db 98 LATLSCLVLLKWNVGDQLQVEGIEFIKSNLELVKDETDQDSLVTDFFIIPSLIREAQS 157
QY 123 LGLDLPYHLPFISIQHOKRQKLOKIPLVNLHNTALLYSLGLEQDVVDWQETINLQSR 182
Db 158 LRLGLPYDLPYIHLQTKRQERLAKLSREEIYAVPSPLLYSLEGIQDIWEVERIMEVQSQ 217
QY 183 DGSFLSSPASTACVPMHTQNKCLHFLNVLFSKFGDYVCHYPLDLPERLWADTVRLG 242
Db 218 DGSFLSSPASTACVPMHTGDAKLEFLNSVMIKFGNFVPCLPVDDLLERLLIIVNIVRLG 277
QY 243 IDRYFKKEIKESLDYVRYWDAERGVMARCNPIPDVDDTAMGLRLTLRLHGVNVSDDL 302
Db 278 IYRHFKEIKALDYVRYHWN-ERGIWGLRNPIADLETTALGFLRLHRYNVSPIAFD 336
QY 303 NFRDEKDPFCFAGQTOIGVTDNLNLYRCSQVCFGEKIMEAKTFTTNHLQNALAKNA 362
Db 337 NFKDANGKFCSTQFNKDVASMLNLYRASQLAFPGENILDEAKSFATKYLREALEKSET 396
QY 363 FDKWAVKDLPGVEYAIKYVWHRSMPLREARSYIEQFGSNDVWLGKTYVYKMLYVSNKY 422
Db 397 SSANNKQNLSEIKYALKTSWHSAPRVEAKRYCQVYRDPYARIKACVYKLPYVNEKF 456
QY 423 LELAKLDPNMVQALHOKETQHIWVNRBSGFNDLFTFRQRPVEMVSVAVSMFEPPEAAC 482
Db 457 LELGKLDNFIIQSIHQEEMKNVTSNFRDGLPLFTFARERPLEFYLVAAGTYEPQYAKC 516
QY 483 RIAYAKTSCLAVIDDLVTHGSLDDLLKFSFAVRWDISVLDSVRDNLKVCFLGLYNT 542
Db 517 RFLFTKVAQLQVLDMDYDTYGLDELKLFTEAVRRWDLSTENLPD-YMKLCYQIYYDI 575
QY 543 VNGFGKDLKEQGRDVLGYLRKRWEGLLASVYKBAEWSAAKYVPTFNEVYENAKVSTALA 602
Db 576 VHEVAWEAKKEQGRDLVFFRKGWEDYLLGYEBAEWLAAYVPTLDEYIKNGITSIGOR 635
QY 603 TVVLNSI-FFTGELLPDYILQOVDL--RSKFLHLVSLTGRINDTKTYQAERNRGLVSS 659

; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/398,395
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: 60/100,993
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/130,628
; PRIOR FILING DATE: 1999-04-22
; PRIOR APPLICATION NUMBER: 60/150,262
; PRIOR FILING DATE: 1999-08-23
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 46
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Abies grandis
US-09-903-012-46

Query Match      51.6%; Score 2050.5; DB 9; Length 782;
Best Local Similarity 50.8%; Pred. No. 4.4e-190;
Matches 379; Conservative 152; Mismatches 206; Indels 9; Gaps 7;

QY 6 GDGET--NPSAYDTAWARIPISIDSGAPQFPQTQWILNNQLPDGSGWGEICIFLAYDRV 53
Db 38 GDGESMITPSAYDTAWARVPAIDGSARPQFPQTQWILNNQLKDGSGWGIQSHFLSDDL 37
QY 64 LNTLACLLTLKWNKGDIOVQKGVFVRKHEMKDEAD-NHRPSGFVFPVPMLEAKS 122
Db 98 LATLSCLVLLKWNVGDQLQVEGIEFIKSNLELVKDETDQDSLVTDFFIIPSLIREAQS 157
QY 123 LGLDLPYHLPFISIQHOKRQKLOKIPLVNLHNTALLYSLGLEQDVVDWQETINLQSR 182
Db 158 LRLGLPYDLPYIHLQTKRQERLAKLSREEIYAVPSPLLYSLEGIQDIWEVERIMEVQSQ 217
QY 183 DGSFLSSPASTACVPMHTQNKCLHFLNVLFSKFGDYVCHYPLDLPERLWADTVRLG 242
Db 218 DGSFLSSPASTACVPMHTGDAKLEFLNSVMIKFGNFVPCLPVDDLLERLLIIVNIVRLG 277
QY 243 IDRYFKKEIKESLDYVRYWDAERGVMARCNPIPDVDDTAMGLRLTLRLHGVNVSDDL 302
Db 278 IYRHFKEIKALDYVRYHWN-ERGIWGLRNPIADLETTALGFLRLHRYNVSPIAFD 336
QY 303 NFRDEKDPFCFAGQTOIGVTDNLNLYRCSQVCFGEKIMEAKTFTTNHLQNALAKNA 362
Db 337 NFKDANGKFCSTQFNKDVASMLNLYRASQLAFPGENILDEAKSFATKYLREALEKSET 396
QY 363 FDKWAVKDLPGVEYAIKYVWHRSMPLREARSYIEQFGSNDVWLGKTYVYKMLYVSNKY 422
Db 397 SSANNKQNLSEIKYALKTSWHSAPRVEAKRYCQVYRDPYARIKACVYKLPYVNEKF 456
QY 423 LELAKLDPNMVQALHOKETQHIWVNRBSGFNDLFTFRQRPVEMVSVAVSMFEPPEAAC 482
Db 457 LELGKLDNFIIQSIHQEEMKNVTSNFRDGLPLFTFARERPLEFYLVAAGTYEPQYAKC 516
QY 483 RIAYAKTSCLAVIDDLVTHGSLDDLLKFSFAVRWDISVLDSVRDNLKVCFLGLYNT 542
Db 517 RFLFTKVAQLQVLDMDYDTYGLDELKLFTEAVRRWDLSTENLPD-YMKLCYQIYYDI 575
QY 543 VNGFGKDLKEQGRDVLGYLRKRWEGLLASVYKBAEWSAAKYVPTFNEVYENAKVSTALA 602
Db 576 VHEVAWEAKKEQGRDLVFFRKGWEDYLLGYEBAEWLAAYVPTLDEYIKNGITSIGOR 635
QY 603 TVVLNSI-FFTGELLPDYILQOVDL--RSKFLHLVSLTGRINDTKTYQAERNRGLVSS 659
QY 660 VQCYMRENPECTEALSHVYGIIDNALKELNLANPASNAPLCVRLLENTFARVMOLF 719
Db 696 IECYKDHPECTEALDHIYSILEPAVKELTREFLKP-DDVPPACKQMLFEETRVTWVI 754
QY 720 YMYRDGFGISDKEMKHVSRTLPDPV 745
Db 755 FKDGDFGVSKLVKDHKECLIEPL 780
```

Db 636 ILLSGVIMDQLLSQALEKVDYPGRRVLTSLNSLISRLADDTKYKAKARGELASS 695  
Qy 660 VOCYMRNPECTEERESHVYGIIDNALKELNWE LANPASNAPLCVRLLFNFAVMOLF 719  
Db 696 IECYMKDHPECTEERELDHIYSILEPAVKELTREFLKP-DDVPFACKQMLFEETRVTWVI 754  
Qy 720 YMYRDGFGISDKEMKDHVSRTLDFDV 745  
Db 755 FKDGGRGVSKLEVNDHIKECLIEPL 780

Search completed: August 17, 2004, 21:12:41  
Job time : 42.5485 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:57:35 ; Search time 14.7449 Seconds  
(without alignments)  
4866.690 Million cell updates/sec

Title: US-10-041-007-39  
Perfect score: 3972  
Sequence: 1 MFQSMGDETPNSAYDTAWV.....GISDKEMKHVSRITLFDPA 746

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320	33.2	802	2 D85035	ent-kaurene synthase
2	1293.5	32.6	823	2 T02959	kaurene synthase A
3	1275.5	32.1	801	2 T06783	ent-kaurene synthase
4	1066.5	25.9	789	2 T09672	ent-kaurene synthase
5	1025.5	25.8	785	2 G96825	hypothetical prote
6	1025.5	25.8	785	2 T52059	ent-kaurene synthase
7	704	17.7	870	2 A96637	hypothetical prote
8	663	16.7	520	2 A56118	vetispiradiene syn
9	660	16.6	599	2 A48863	limonene cyclase -
10	654.5	16.5	600	2 F71434	probable limonene
11	650.5	16.4	1024	2 G71434	probable limonene
12	647.5	16.3	554	2 S68365	(+)-delta-cadinene
13	643.5	16.2	554	2 S68366	(+)-delta-cadinene
14	633.5	15.9	550	2 T03714	5-epi-aristolochene
15	633.5	15.9	559	2 T08174	sesquiterpene cycl
16	633	15.9	591	2 H84633	probable limonene
17	565	14.2	548	2 T06266	germacrene C synth
18	559	14.1	548	2 T06265	germacrene C synth
19	469.5	11.8	632	2 E96723	hypothetical prote
20	460.5	11.6	598	2 H86460	hypothetical prote
21	448	11.3	350	2 C56118	vetispiradiene syn
22	443	11.2	608	2 G86443	probable terpene s
23	416.5	10.5	573	2 T05328	hypothetical prote
24	403.5	10.2	582	2 C71424	hypothetical prote
25	393	9.9	530	2 G96588	hypothetical prote
26	389.5	9.8	535	2 T06285	hypothetical prote
27	388.5	9.8	421	2 C96642	hypothetical prote
28	375	9.4	598	2 T05329	hypothetical prote
29	374	9.4	598	2 T00509	probable vetispira

30	371	9.3	501	2 H70972	probable cyclase -
31	369	9.3	598	2 F96684	probable terpene s
32	351.5	8.8	300	2 B56118	vetispiradiene syn
33	349.5	8.8	612	2 T05331	hypothetical prote
34	330.5	8.3	471	2 T06287	hypothetical prote
35	313	7.9	383	2 D71424	hypothetical prote
36	271	6.8	952	2 JC7227	ent-kaurene synthase
37	248	6.2	946	2 T00024	ent-kaurene synthase
38	219.5	5.5	516	2 T10874	y4kt protein - Rhi
39	193.5	4.9	516	2 I40214	hypothetical prote
40	166	4.2	203	2 H96525	probable terpene c
41	151.5	3.8	148	2 T03982	5-epi-aristolochene
42	134	3.4	1131	2 T30951	hypothetical prote
43	134	3.4	3864	2 D87757	protein C44B4.1a [
44	122.5	3.1	757	2 C70034	conserved hypothe
45	122.5	3.1	1274	2 S28279	hypothetical prote

ALIGNMENTS

RESULT 1

D85035

ent-kaurene synthetase A-like protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001

C:Accession: D85035

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: AB5001; MUID:20083488; PMID:10617198

A:Accession: D85035

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-802 <STO>

A:Cross-references: GB:NC\_001268; NID:G7269763; PIDN:CAB77763.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4G02780

A:Map position: 4

Query Match 33.2%; Score 1320; DB 2; Length 802;  
Best Local Similarity 36.6%; Pred. No. 2e-84;  
Matches 281; Conservative 144; Mismatches 252; Indels 90; Gaps 17;

QY	1	MFQSMGDETPNSAYDTAWVARI	PSIDSGAGPQPTQLWILNNQLPDGSGWCECIFY	60
DB	104	ILRLNTDGEITTSAYDTAWVALIDA	--GDKTAPFSAVKWIAENQLSDGSGWDAYLPSYH	161
QY	61	DRVLTNLTACLLTLKWNKGD	IQVQGVFVRKHEMKDEADNHRFSGFVFPVFPAMLDEA	120
DB	162	DLRLNTLACVVALRSWNLFP	HCNKGITFFRENIKLEDENDHMPIGFEVAFPSLLEIA	221
QY	121	KSLGLDLPVHLFPISQIHOK	RQKLOKIPLVNLHNTALLYSLLEGQDVVDVQOETNLQ	180
DB	222	RGINDIVPVDSPVLKDIYAK	ELKLTIPRIPMHKIPPTLLHLEGWRD-LDWEKLKQJ	280
QY	181	SRDGSFLSSPASTACVFMHT	QNKRCILHFLNVLSEKFGDVVPCPHPLDLFERLWAVTVER	240
DB	281	SDGSGFLFSPSTAFAPMQTR	DSNCLYLRLNAVKRFGVGVNVPVDLFEHIVDLQR	340
QY	241	LGIIDRYFKKEIKESLDYV	RYWDAERGVGWARCNPIPDVDDTAMGLRILHLHGVNSDV	300
DB	341	LGISRYFEEIKECIDYVHYR	W--TDNGICWARGSHVQDIDDTAMAFLLRHQGVQVSADV	399
QY	301	LSNERDEKGFCEFCAGTOIG	VTNLTNLVRCQVCPGKIMEAKTFTTNHLQNALAKN	360
DB	400	FXKNP-EKEGFFCFVQGSNO	AVTGMNLYRASOLAPRBEILKNAKEFSTNYLLEKERE	458
QY	361	NAPDKWAVKOLPGEVEYAI	KYPWHRSMPLRLARSYIEQF-GNSDVMLGTVTKMLYVSN	419
DB	459	ELIDKWIIMKDLFGEIGFAL	EIPWYASLPRVETRFYIDQYGGENDVWIGTKLRMPVNN	518
QY	420	EKYLELAKLDFTNNVQALH	QKQTKQHVSWWRSGFNDLTFTTRQRPVEMYFSAVSMFPEF	479

Db 519 NGYLELAKQDYNCOAQHLEWDIFQKWEENRLESEWGRVRSLECELYLAATIFESER 578  
Qy 480 AACRIAYAKTSCALVILDDLYDTHGSLDLDLKFSEA-----VRRWDISVLD----- 525  
Db 579 SHERMWAKSVLVKAISSSFGF--SSDSRSFSQFHEFYIANARRSDHFFHFNDRNRLDR 636  
Qy 526 --SVYRDNLKVCFLGLYNTVNGFGDGLKEQGRDVLGYLRKWEGLLASITYKEASWSAAK 583  
Db 637 PGSVOASRLAGVLIG--TLNQMSFDLPMSHGRDYNLLYSW-----GDM----- 679  
Qy 584 YVPTFNEVENAKVSIALATVVL--NSI--FFTCELLPDYILOQVDRLSKFLHVLSTG 638  
Db 680 -MEKWLXGDEGEGLAVKVMILMKNDLTNEFT-----HTHFVRLAEIIN 724  
Qy 639 RLINDTKTYQAEARNRGELVSSVQCYMRENPECTEBEALSHVYGIIDNALKEINWELANPA 698  
Db 725 RICLPQRYLKARN-----DEKBTIKSMEKMGKMWELALSSED----- 764  
Qy 699 SNAPLCVRLLFNTARVMQVFMVYRDGFGISDKEMKHVSRTLPDPV 745  
Db 765 -----TFRDSVITFLDVAKAFY-----FALCGDHLQTHISKVLFOKV 802

RESULT 2  
T02959  
kaurene synthase A - maize  
C:Species: Zea mays (maize)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 29-Oct-1999  
C:Accession: T02959  
R:Bensen, R.J.; Johal, G.S.; Crane, V.C.; Tossberg, J.T.; Schnable, P.S.; Meeley, R.B.;  
Plant Cell 7, 75-84, 1995  
A:Title: Cloning and characterization of the maize An1 gene.  
A:Reference number: Z14794; PMID:7696880  
A:Accession: T02959  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-823 <BEN>  
A:Cross-references: EMBL:L37750; NID:G576885; PIDN:AAA73960.1; PID:G576886  
C:Genetics:  
A:Gene: AN1  
C:Function:  
A:Description: cyclase involved in gibberellic acid biosynthesis

Query Match 32.6%; Score 1293.5; DB 2; Length 823;  
Best Local Similarity 36.5%; Pred. No. 1.5e-82;  
Matches 289; Conservative 146; Mismatches 243; Indels 113; Gaps 24;

Qy 1 MFQSGDGETNPSAYDTAWARIPSI--DGSGAPQFPOTLOWILNQLPDGSGWGEICPL 58  
Db 100 MLRSNWDGTSASAYDTAWAMVPKVGDDGGQAPQFPATRVIVDQLEPDGSGWDSALPS 159  
Qy 59 AYDRVLNTLACLLTLKWNKGDIQVQGVFVRKHEEMKDEADNRHPSGFEVFPFAMLD 118  
Db 160 AYDRMINTLACVVALTKWLEPARCEAGLSFLHNNWRLAEAEASMPFGIEAFPSLIQ 219  
Qy 119 EAKSLG-LDLPHLFPISQIHQKQKQKLOKIPNLVNLHQTALLYSLGLEQVDVWQETIT 177  
Db 220 TARDLGVDVFPFGHEPALOSIYANREVKLRIPRDMHVRVPTSLHSLEGMPP-LDMPRL 278  
Qy 178 NLQSRDGSFLSSPASTACVFMHTONKRLHFLNVLKSGDVVPCHYPLDLFLERLWAVDT 237  
Db 279 NLQSCDGSFLSPSATAVALMOTGKKCFEYIDRIVKKFGVGNVPVYFDLFEHLWVDR 338  
Qy 238 VERLGDIRYFKKIKESLDYVRYNDAGRGVWARGCNPIPDVDDTAMGLRLILRHGYNYS 297  
Db 339 LERLIGISRYFIEIQECMDYVRHW--TEDGICWARKSNVKVDVDTAMAFELRLHGYNS 397  
Qy 298 SDVLENFRDEKDFCFAGQTOIGTDNLNLRCQVCFPGKIMEEAKTFTHNLQNAL 357  
Db 398 PSVFNKF-EKDGEEFCFVGQSTQAVTGMYNLNRASQISFGEDVLHRAVFSYFLRQRE 456  
Qy 358 AKNNAFDKWA VKDLPGVEVEYAIKVPWHRSMRLEARSIEQF-GSNDVWLKGVTKVYKMLY 416

## RESULT 3

T06783

ent-kaurene synthase A (EC 2.5.1.-) - garden pea

C:Species: Pisum sativum (garden pea)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000

C:Accession: T06783

R:Ait-Ali, T.; Swain, S.M.; Reid, J.B.; Sun, T.P.; Kamiya, Y.

Plant J. 11, 443-454, 1997

A:Title: The LS locus of pea encodes the gibberellin biosynthesis enzyme

A:Reference number: Z15810; PMID:97260958; PMID:9107034

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-801 &lt;AIT&gt;

A:Cross-references: EMBL:U63652; NID:G2160543; PIDN:AAB58822.1; PID:G2160544

A:Experimental source: cultivar Toradag

C:Genetics:

A:Gene: LS

C:Function:

A:Description: converts the geranylgeranyl diphosphate into copalyl diphosphate

A:Pathway: the first step of the gibberellin biosynthesis

C:Keywords: transferase

Db 276 OSQGSFLPSSSTAFALMOTKQGNCLKYLNNVVKFKNGGVNVPVDLFEHIIWVDRLE 335  
Qy 240 RLGIDRYPKGIKESLDVYRYWDAGRGVGMARCNPIPDVDDTAMGLRILRLHGVNVSDD 299  
Db 336 RLGSRPPEHIEKDCMNVSKIW-SEKIGCWARNVQDIDDTAMAFRLRLHGHQVSAH 394  
Qy 300 VLENPRDEKGPFCFAGQTOGVTDNLNLYRCQVCPGEGKIMEBAKPTTNHLQNALAK 359  
Db 395 VFKEHF-ERNGBFFCFAGQTOAVTCGMYNLFRASQVLPFGKEKLEHAKHPSAKVLKKEBA 453  
Qy 360 NNAFDKWAQKDLQVEYAIKYPWHRSPRLARSYIEQFGS-NDVWLGKTVYKMLYVS 418  
Db 454 NELIDKWIIMKNLPEVGYALDMPWYANDRIETRFYIDQYGAESDVMIGKTLRYMAYVN 313  
Qy 419 NEKYLELAKLDFNMVQALHOKETQHIVSWMRSGFNDLFTFRQRPVEMYFVSVAVMFPE 478  
Db 514 NNNYLELAKLDYNNCAQHLEIWNVIQTYWLESRLGEPGLSKRDLLAYFLATGSIPEPE 573  
Qy 479 FAACRIAYAKTSCLAVIL-----DDLVDYTHGSLDDLKLPSEAVRRWDISVLDSDVDLQ 532  
Db 574 RSHERLAWAKTTALLETIKCYVRNEDL-----RKDFAKKFNHDHIDVRDYSIARRMKRNT 528  
Qy 533 KVCFPL-GLYNTVNGFGDKLKEQGRDVLGY-LRKVWEGLLASYTKEAWSAAKYVPTFNE 590  
Db 629 EHELVEISLFAIGELISVDRLSYGHEI-GYDWHQCWKWLSWSQSEGDKCEG-----580  
Qy 591 YVENAKVSIATATVNLVSIFFTGELLDPYIIQQVDLRSKFLHLVSLTGLRLINDTKYQAE 550  
Db 681 ----AELLIIQINLCSN-----HWISGSPSQSTQIHLQLTNSICHKLSYQKD 726  
Qy 651 RNRGELVSSVOCYKRENPECTEERALSHPYGIIDNALKELWELANPASNAPLCVRRLLP 710  
Db 727 KE-----LKGISC--QEN--ITNSEVESKQELVQMFQK-----CPNDIDF 764  
Qy 711 NTAIRMQLFTMYRDFGSGIS-----DKEMKDHVSRTLFDPV 745  
Db 765 N----VKMTFTTIAKSYFYAFCDSKRTINFHIAKLVLEK 800  
RESULT 4  
T09672  
ent-kaurene synthase B (EC 2.5.1.1) - winter squash  
C/Species: Cucurbita maxima (winter squash)  
C/Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
R/Yamaguchi, S.; Saito, T.; Abe, H.; Yamane, H.; Murofushi, N.; Kamiya, Y.  
Plant J. 10, 203-213, 1996  
A/Title: Molecular cloning and characterization of a cDNA encoding the gibberellin biosyn-  
A/Reference number: 216814; MUID:96367664; PMID:8771778  
A/Accession: T09672  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-789 <YAM>  
A/Cross-references: EMBL:U43904; NID:gl431869; PIDN:AAB39482.1; PID:gl431870  
A/Experimental source: immature seeds  
C/Function:  
A/Description: catalyzes the conversion of copalyl diphosphate to ent-kaurene  
A/Pathway: Gibberellin biosynthesis  
A/Note: terpene cyclase  
C/Keywords: transferase  
Query Match 26.9%; Score 1066.5; DB 2; Length 789;  
Best Local Similarity 33.4%; Pred. No. 1.2e-66;  
Matches 244; Conservative 151; Mismatches 302; Indels 33; Gaps 16;  
Qy 9 ETNPSAYDTAWARIPSIDGSGAQFPQTLOWILNNQLPDGSGW-----EGCIFLAYDRVL 54  
Db 56 ELSYSDYDTAWAVWSPSNLNLQFLFPFCINVLDSQHADGSGWGLLHNDQLLMKA--NLL 113  
Qy 65 NTLACLTLTKLWNGDIOVQGVFVRKHMEEMKDEADNHRPSPGVVFPAMLDKASIG 124  
Db 114 STLACVLTKKWNIGHDHMSKALDFIKSNIASATDE--NQRSPVGFDIIFPGMIEYAKOLN 172

Qy 125 LQLPYHLPIFSQIHQKROKKLQKIPNLVNLHNHOTALLYSLLEGLOQVVDVWQEIETNLQSRDG 184  
Db 173 LNLPLAPTVDVALVRKKELELRSCRN-SEGGKAYLAYVSEGIQKLDQWDMVQYQKNG 231  
Qy 185 SFLSSPASTACVPMHTQNKRLHPLNFVLSKSGDYVPCYHPLDLPERLWAVDTVERLGID 244  
Db 232 SYFNGSPSTTAAAFMHRNDGCFDYLRSLLQKPGSGVPTIYPLDIYARLHWVDSLQKFGIA 291  
Qy 245 RYFKKEIKESLDVYRYW-DAERGUGWARCNPIPDVDDTAMGLRILRLHGVNYSDDVLEN 303  
Db 292 RHFKKEIRSVLDETYRCWMOGEE-----NIFLDASTCAMAFLMLRVEGYDVSSDQLTQ 344  
Qy 304 FRDEKGDFF--CFAGQTO-IGVTDNLNLYRCQ-VCPGEGKIMEBAKPTTNHLQNALAK 359  
Db 345 FSE---DIPFNCLGGYKDFGA--SLELYKASQIIITHPDESVLNINSMWTSRFLKHGLSS 399  
Qy 360 NNAFDKWAQKDL-LPGEVEYAIKYPWHRSPRLARSYIEQFGSNDVWLGKTVYKMLYV 417  
Db 400 DSV---WSRDTOSVVKQEAVALDEFFYNATLERLISKRAMESYSGDILVRSKSPYACINF 456  
Qy 418 SNEKYLELAKLDFNMVQALHOKETQHIVSWMRSGFNDLFTFRQRPVEMYFVSVAVMFPE 477  
Db 457 GHQDFLELAVEDFNTLQRIHLKELBELQWVVENKLDLKFRLHLGYCYFAAAATLTD 516  
Qy 478 EFAACRIAYAKTSCLAVILDDLVDYTHGSLDDLKLPSEAVRRWDISVLDSDVDLQKVCFL 537  
Db 517 ELHDARIAWAQNGVLTITVVDDFYDGGGSEBELNLLIHELVEKMDPDGEGYCSKDVEIVFL 576  
Qy 538 GLYNTVNGFGDKLKEQGRDVLGYLRKVWEGLLASYTKEAWSAAKYVPTFNEVENAKV 597  
Db 577 ALHSTVCEIGRALVWQGRSVNRNIDGWLALLKVMRKEAESTNKNVPSMGETMEQAHV 636  
Qy 598 SIATATVNLVSIFFTGELLDPYIIQQVDLRSKFLHLVSLTGLRLINDTKYQAEENRGELV 657  
Db 637 SPALGPILPLMFLFFVGPKLSEEMIGSCYQ-KLYKLMSTAGRLKNIRSDRECKEKL- 694  
Qy 658 SSVQCYMRB-NPECTEERALSHPYGIIDNALKELWELANPASNAPLCVRRLLFNTRVM 716  
Db 695 NILSLWMIDGGNGVTKEAIEAKGDFERAIRELLGLVLQENTTIPRACKDLFWKLSIV 754  
Qy 717 QLFYMYRQGF 726  
Db 755 NLFYMEDDGY 764  
RESULT 5  
G96825  
hypothetical protein T8K14.12 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: G96825  
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: G96825  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-785 <STO>  
A/Cross-references: GB:AE005173; NID:g4835764; PIDN:AAD30231.1; GSPDB:GN00141  
C/Genetics:  
A/Gene: T8K14.12  
A/Map position: 1  
Query Match 25.8%; Score 1025.5; DB 2; Length 785;  
Best Local Similarity 33.0%; Pred. No. 8.7e-64;





A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:210161719; PMID:11130712

A:Accession: A96637

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-870 <STO>

A:Cross-references: GB:AE005173; NID:G2443889; PIDN:AAB71482.1; GSPDB:GN00141

C:Genetics:

A:Gene: Flp17.15

A:Map position: 1

Query Match 17.7%; Score 704; DB 2; Length 870;  
Best Local Similarity 27.5%; Pred. No. 3.6e-41;  
Matches 219; Conservative 157; Mismatches 301; Indels 120; Gaps 30;

QY 11 NPSAYDTAWARIPS-----IDSGAPQFPQTILQILNNO-LPDGSGW----- 31

DB 37 SPSAYDTAWLSMBEDINVDNDELKPMFGCLDWIMCNQAREGFWMNSTSYTTVADGRD 36

QY 52 --GER--CIPLAYDRVNTLACLLTKIWNKGDIOVQKGVFVRKHMEKDEADNHR-- 105

DB 97 EDGEKMCII-----LTSLACVVALQKWNIGCFHLHKGTRYIERNTEMIIGKIYNEGS 150

QY 106 -PSGFEVVPAPMLDEAKSILGLDPYH---LPTSIHQKQKLOKIPLVNLHNTALL 161

DB 151 YPRWFAIKFTGILELAQKLGHLHFVSSRCIEMIKGFYQRBIIQREKLVHDCNKPLLA 210

QY 162 YSLEGL-----QVVVDWOBETNLQSRDGSFLSSPASTACVPMHTQNKRCILHFLNVL 212

DB 211 Y-LEVLPSKLYVYVQEDII-----VKSLSMDGSLFQSPSATASAFMLTENTKCLAYLQNL 265

QY 213 LSKFGDYVPCHYPL--DLFERLWADTVRLGIDRYFKKEIKESLDYVY--WDAERG 267

DB 266 VQKCPNGVQKPYPLMEDLI-KLSMVNLTESTGLGFEFFGIEIEHVLEQVYRYEEKDFER- 323

QY 268 VGMARCNPIPDVD---TAMGLRLRLHGVNVS-----SDVLENFRDEKDFPCF 314

DB 324 -----MPMSYLADQLHKDSLAFLRMHGRDVSPPSPFCWFLNDQETRNHLERNIDSFL 377

QY 315 AGQTQIGVTNINLYRCQVCPFGPKIMEEAKTFTTNLQNALAKNNAFDMWAKKOLP- 373

DB 378 V-----ILSVTRATDLMPEGHEIDQAREYT-----RNLLKERSIK-----EKWPI 420

QY 374 --GEVEYAKYPMHRSMPRLARSYIEQFGSNVLMGKTYYKMLYVS-NEKYLELAKLDF 430

DB 421 YLNKIMHELSTPWIALRKLHDMWIEDKNVLSMERKASFLRLHSSYSKLTLLHAAFN 480

QY 431 NMVQALHOKETQHIHVSWRRESGFNDLITTRQRPVEMYFSVASM-PEPEFAACRIAYAKT 489

DB 481 EFQQAQYCRELEEL-----TIEKTTYCYFATVTSPLPYEYAIKFGKLA-AKT 525

QY 490 SCLAVILDDLYTHGSLDDKLFSAVRRWDISVLDSVRDNQKVCFLGLYNTVNGFGKD 549

DB 526 AILITIADDFDEKGSFNDLEGLTKAVLRWEEELKSYGN-----IIFRALDIVRETANT 581

QY 550 GLKSGQRDLVGLYRKVWEGLLASYYTKEAFAWSAAKYVPTFNEYENAKVSIALATVVLNSI 509

DB 582 CRTHKTDIIIVHLRNIGWETPESWLRREAESWKGGTSSMDVEYIRNGMISIAAHTIALSIS 541

QY 610 FFTGELLDPYILQOVDLSKFLHLVSLTGRLLNDTKYQAEERNRGLYSSVQCYRENPE 569

DB 642 CLMEPCFPHNKLKPNYDS-ITTLMIIPRLNDLQSQYKQEQGKM-NSVLLHMKNHFG 599

QY 670 CTEEALSHVYGIIONALKE-LNWEILANPASNAPLCVRRLLFNTARVMOLFMYMRDGGFI 728

DB 700 LEIEDSIAHEIKLIDSKRKEFLHVLVDGLSLDPKPKCEIHMSSCCVPEMFNKKNRDY- 758

QY 729 SDKEMKDHVSRTLPDPV 745

DB 759 SNTMLDHIKKALYDPI 775

RESULT 8

A56118

vetispiradiene synthase 1 - Hyoscyamus muticus (fragment)

C:Species: Hyoscyamus muticus

C:Date: 19-Oct-1995 #sequence\_revision 19-Oct-1995 #text\_change 29-Sep-1999

C:Accession: A56118

R:Back, K.; Chappell, J.

J. Biol. Chem. 270, 7375-7381, 1995

A:Title: Cloning and bacterial expression of a sesquiterpene cyclase from Hyoscyamus muticus

A:Reference number: A56118; MUID:95221394; PMID:7706281

A:Accession: A56118

A:Molecule type: mRNA

A:Status: preliminary

A:Residues: 1-520 <BAC>

A:Cross-references: GB:U02188; NID:G763422; PIDN:AAA86337.1; PID:G763423

C:Superfamily: vetispiradiene synthase 1

Query Match 16.7%; Score 663; DB 2; Length 520;

Best Local Similarity 27.2%; Pred. No. 1.2e-38;

Matches 157; Conservative 130; Mismatches 216; Indels 74; Gaps 12;

QY 174 QBITNLQSRDGSFLSSPASTACVPMHTQNKRCILHFLNVLKSGFDYVPCHYPLDLPERLW 233

DB 11 QBITLKEQSTWLSAACGTT-----LTKELN 37

QY 234 AVDTVERLIGIDRYFKKEIKESLDYVYR---YWDARGVGVHARCNPDPDVTDDTAMGLRLR 290

DB 38 LIDIIERLGIAYHFKEQIEDMLDHIYRADPYFA-----HEYNDLNTSSVQFRLLR 88

QY 291 LHGVNVSVDLENFRDEKDFCFACQTOIGVTDNLNLYRCQVCPFGPKIMEEAKTFTT 350

DB 89 QHGVNVSPIFERFQDANGK---FKESLRSDINGLNLNLYEASHVTRTKEDILEALVFSV 145

QY 351 NHLONALAKNNAFDMWAKKOLPGEVEYAIKYPHRSMPRLARSYIEQFGSNVLMGKT 410

DB 146 GHLESAAPH-----LKSPLSKQVTHALEQSLHKSIPRVEIRYPI-----S 185

QY 411 VTKMLYVSEKYLELAKLDFNMVQALHOKETQHIHVSWRRESGP-NDLITTRQRPVEMYFS 469

DB 186 IYEEEFKNDLLRFALKDYNLQMLHKEHSEVSRWMDKLDLDFVTLTPYADRARECYFW 245

QY 470 VAVSMPEPEFAACRIAYAKTSCLAVIDLDLYTHGSLDDKLFSVARRVDISVLDSVRD 529

DB 246 TWGVTAEPQYSQARVLMKTIAMISIVDDTDAIGYVKELEVYDIAQRWDISQIDRLPE 305

QY 530 NOLKVCFLGLYNTVNGFGDKLKEQGRDLVGLYRKVWEGLLASYYTKEAFAWSAAKYVPTFN 589

DB 306 -YMKISYKALLDLYDDYEKELSKDGRSDVVHYAKERMEKEIVGVNFIEGKWFIEGYMFSVS 364

QY 590 EYVENAKVSIALATVVLNSIIFPTGELLDPYILQOVDLSKFLHLVSLTGRLLNDTKTYQA 649

DB 365 EYLSNALATSTYLLTTTTSYLGKMSATKSH-FEWLATNPRILEANATLCRWVDDIATYEV 423

QY 650 EERNRGLYSSVQCYRENPECTEEALSHVYGIIONALKELNWEILANPASNAPLCVRRLL 709

DB 424 ENRGQIATGIECYMDYGVST-EVAMEKPEQENADIAWKDVBNEILRTPVPSSEILTRIL 482

QY 710 FNTARVMOLFMY-RDGFGISDKEMKDHVSRTLPDPV 745

DB 483 -NLARIIDVYKHNQDGYTHPEKVLKPHIALVVDISI 518

RESULT 9

A48863

limonene cyclase - spearmint

C:Species: Mentha spicata (spearmint)

C:Date: 12-May-1995 #sequence\_revision 19-May-1995 #text\_change 05-Nov-1999

C:Accession: A48863

R:Colby, S.M.; Alonso, W.R.; Katahira, B.J.; McGarvey, D.J.; Croteau, R.

J. Biol. Chem. 268, 23016-23024, 1993

A:Title: 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA is

A:Reference number: A48863; MUID:94043077; PMID:8226816

A:Accession: A48863

A:Status: preliminary



Db 27 LLEQLELDTLQRLGVSYHPEQEI KKTLTNNVKNVRAHKNRDNRR---WG-----DL 77  
Qy 280 DDTAMGLRIILRHGY---NVSSDVLNFRDPSKGFPCFAGTQIGVTDNLNLYRCSQVCF 336  
Db 78 YATALEFRLLRQHDVDFGNI GVDL--DDKOIKG-----ILSLYRASYLST 120  
Qy 337 PGEKIMEAKTFTTNHLQNALAK-NAPDKWAVKOLPGEVEYAIKYPWHRSMPLRLARS 395  
Db 121 RIDTKLESIIYITTKRLKPFVEVKNETKSYTLRR---MVIHALEMPYHRRVGLRLARW 176  
Qy 396 YIEQFSGNDVNLGTVKTVKMLVSNKYLELAKLDFNMVQALHOKETOHIHVSWRESGF-N 454  
Db 177 YIEVYGER-----HDMNPILLELAKLDFNFVQAIHQDELKSLSSWWSKTGLTK 224  
Qy 455 DLFTTRQRPVEMYSVAVSMPPEPFAACRIAYAKTSCIAVLDDLDYTHGSLDDKLFSF 514  
Db 225 HLDVPRDRITEGYSSVGVMPYEPFAVHRQMLTKVFMILITIDDIYDYGTLLEELQFTT 284  
Qy 515 AVRWDIVLSDVRDNLKVCFLGLYNTVNGFGDKLKEQGRDVLGYLRKVEGLLASVT 574  
Db 285 IVEKWDVNRLEEL-PNMYKLCFLCLVNEINQIGFVLKDRGNFVPIYLKESWADNCTTFL 343  
Qy 575 KEAEWSAAKVTPTNEVYENAKVSIALATVVLNSITFTGELLPDYILQQVDLRSKFLHLV 534  
Db 344 KEAKWYSKYKPNFEYMQNWISSVPTILH-LFC---LLSD---QTLILGSYNHSV 396  
Qy 635 SLTG-----RLINDT-----KTYQABR-NRGELVSSVQCYMRENPECTEERALSHTV 680  
Db 397 VRSSATILRLANDLATSVSHTGTTNTTEELARGDTMKSVCQHMHT-GASEAESRAVIQ 455  
Qy 681 GIIDNALKEINWELANPASNAPLCVRRLLP-----NTARVMQLFYMYRDGFGISDK-EMK 734  
Db 456 GIYGVAVDDLLNMEKKS-----CRHQGFLEAAANLGRVACQVYQYDGHGCPDKAKTV 508  
Qy 735 DHVSR 739  
Db 509 NHVSR 513

RESULT 12  
S68365  
(+)-delta-cadinene synthase isozyme XC1 - Gossypium arboreum  
C:Species: Gossypium arboreum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: S68365  
R:Chen, X.Y.; Chen, Y.; Heinstein, P.; Davison, V.J.  
A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a cat  
A:Reference number: S68365; MUID:96132653; PMID:8554317  
A:Accession: S68365  
A:Molecule type: mRNA  
A:Residues: 1-554 <CHE>  
A:Cross-references: EMBL:U23206; NID:g1045311; PIDN:AAA93064.1; PID:g1045312  
A:Experimental source: cultivar Nanking  
C:Superfamily: vetispiradiene synthase 1  
C:Keywords: phytoalexin biosynthesis

Query Match 16.3%; Score 647.5; DB 2; Length 554;  
Best Local Similarity 30.9%; Pred. No. 1.7e-37;  
Matches 162; Conservative 103; Mismatches 210; Indels 49; Gaps 12;

Qy 230 ERLWAVDTVERLIGIDRYPKKIKESLDYVY-RYWDARGVGWARNCPIDVDVDTAMGLRI 288  
Db 72 QKLAFLDSVQRLGVSYHPTKEIEDELENIYHNNDNEN-----DLYTTSIRFL 120  
Qy 289 LRLHGVNVSVDLENFRDEKDFPCFAGTQIGVTDN-----LNLYRCSQVCFPGEKIMEE 344  
Db 121 LREHGVNSCDVFNKFKDEQGNF-----KSSVTSVDRGLLELYQASYLRVHGEDILDE 173  
Qy 345 AKTFTTNHLQNALAKNAPDKWAVKOLPGEVEYAIKYPWHRSMPLRLARSYIEQFSGND 404  
Db 174 AISFTTHLSLAVA-----SLDHPLESEVSHALKQSIRGLPRVREARHYL----- 218  
Qy 405 VWLGTVYKMLVSNKYLELAKLDFNMVQALHOKETOHIHVSWRESGF-NDLTFTTRQP 463  
Db 219 -----SVQODIESHNKVLLEFAKIDFNMQVQLHRKELSELRSRWKOLDPORKLPYADR 273  
Qy 464 VMYFSVAVSMPPEPFAACRIAYAKTSCIAVLDDLDYTHGSLDDKLFSFVARRWDISV 523  
Db 274 VEGYFWISGVYEPQYSLGRKMLTKVIAMASIVDDYDYSATYVEELIPYTKAERWDIKC 333  
Qy 524 LDSVRDNLKVCFLGLYNTVNGFGDKLKEQGRDVLGYLRKVEGLLASYTKAEWSAAK 583  
Db 334 IDELPE-YMKPSYKALLDYYEEM-QVLAHGRQYRVEYAKNAMIRLAQSYLVEAKWTQ 392  
Qy 584 YPTFNEVYENAKVSIALATVVLNSITFTGELLPDYILQQVDLRSKFLHLVSLTGLIN 643

Qy 405 VWLGTVYKMLVSNKYLELAKLDFNMVQALHOKETOHIHVSWRESGF-NDLTFTTRQP 463  
Db 219 -----SVQODIESHNKVLLEFAKIDFNMQVQLHRKELSELRSRWKOLDPORKLPYADR 273  
Qy 464 VMYFSVAVSMPPEPFAACRIAYAKTSCIAVLDDLDYTHGSLDDKLFSFVARRWDISV 523  
Db 274 VEGYFWISGVYEPQYSLGRKMLTKVIAMASIVDDYDYSATYVEELIPYTKAERWDIKC 333  
Qy 524 LDSVRDNLKVCFLGLYNTVNGFGDKLKEQGRDVLGYLRKVEGLLASYTKAEWSAAK 582  
Db 334 IDELPE-YMKPSYKALLDYYEEM-QVLAHGRQYRVEYAKNAMIRLAQSYLVEAKWTQ 391  
Qy 584 YPTFNEVYENAKVSIALATVVLNSITFTGELLPDYILQQVDLRSKFLHLVSLTGLIN 642  
Db 332 NYKPSFEFKAALPTCGYAMLAITSFVGMDIVTPTETFKAAASDPKIIQASTIICRFMD 451  
Qy 643 DTKTYQABRNRGELVSSVQCYMRENPECTEERALSHTVGIIDNALKEINWELANPASNAP 702  
Db 452 DVAEHKFKHRRDDCSAIECYMEY-GVTAQRAYDVFNKHVESAWKDLNQEFLLKP-TEMP 509  
Qy 703 LCVRRLLNTARVMQLFYMYRDGFGISDKEMKHDSRTLPDPVA 746  
Db 510 TEVLNRSNLARVMDVLYREGDGYTVYGAAGGITSLLEPIA 553

RESULT 13  
S68366  
(+)-delta-cadinene synthase isozyme XC14 - Gossypium arboreum  
C:Species: Gossypium arboreum  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 29-Sep-1999  
C:Accession: S68366  
R:Chen, X.Y.; Chen, Y.; Heinstein, P.; Davison, V.J.  
A:Title: Cloning, expression, and characterization of (+)-delta-cadinene synthase: a cat  
A:Reference number: S68365; MUID:96132653; PMID:8554317  
A:Accession: S68366  
A:Molecule type: mRNA  
A:Residues: 1-554 <CHE>  
A:Cross-references: EMBL:U23205; NID:g1045313; PIDN:AAA93065.1; PID:g1045314  
A:Experimental source: cultivar Nanking  
C:Superfamily: vetispiradiene synthase 1  
C:Keywords: phytoalexin biosynthesis

Query Match 16.2%; Score 643.5; DB 2; Length 554;  
Best Local Similarity 30.8%; Pred. No. 3.2e-37;  
Matches 161; Conservative 103; Mismatches 212; Indels 47; Gaps 10;

Qy 230 ERLWAVDTVERLIGIDRYPKKIKESLDYVY-RYWDARGVGWARNCPIDVDVDTAMGLRI 288  
Db 72 QKLAFLDSVQRLGVSYHPTKEIEDELENIYHNNDNEN-----DLYTTSIRFL 120  
Qy 289 LRLHGVNVSVDLENFRDEKDFPCFAGTQIGVTDN-----LNLYRCSQVCFPGEKIMEE 344  
Db 121 LREHGVNSCDVFNKFKDEQGNF-----KSSVTSVDRGLLELYQASYLRVHGEDILDE 173  
Qy 345 AKTFTTNHLQNALAKNAPDKWAVKOLPGEVEYAIKYPWHRSMPLRLARSYIEQFSGND 404  
Db 174 AISFTTHLSLAVA-----SLDYPLESEVSHALKQSIRGLPRVREARHYL----- 218  
Qy 405 VWLGTVYKMLVSNKYLELAKLDFNMVQALHOKETOHIHVSWRESGF-NDLTFTTRQP 463  
Db 219 -----SVQODIESHNKVLLEFAKIDFNMQVQLHRKELSELRSRWKOLDPORKLPYADR 273  
Qy 464 VMYFSVAVSMPPEPFAACRIAYAKTSCIAVLDDLDYTHGSLDDKLFSFVARRWDISV 523  
Db 274 VEGYFWISGVYEPQYSLGRKMLTKVIAMASIVDDYDYSATYVEELIPYTKAERWDIKC 333  
Qy 524 LDSVRDNLKVCFLGLYNTVNGFGDKLKEQGRDVLGYLRKVEGLLASYTKAEWSAAK 583  
Db 334 IDELPE-YMKPSYKALLDYYEEM-QVLAHGRQYRVEYAKNAMIRLAQSYLVEARWTLQ 392  
Qy 584 YPTFNEVYENAKVSIALATVVLNSITFTGELLPDYILQQVDLRSKFLHLVSLTGLIN 643

Db 393 YKSPFEEFKANALPTCGYAMLAIITSVFGMGDIIVTPTETFKWAANDPKIIQAOSTIICREWDD 452

Qy 644 TKYQAEARNRGELVSVQCYRENPECTEERALSHPYGIIDNALKELNWLANPASNAPL 703

Db 453 VAHKFKHRRDDCSAIECYMEY-GVTAQEAAYOVFNKHVESAWKVNKEFLKP-TEMPT 510

Qy 704 CVRRLLFNFTARVMOLFMYNRDQFGISDKEMKHVSRTLPDPVA 746

Db 511 EVLNRSNLARVMDVLYREGDGYTVGKAAGGITSLLIEPVA 553

RESULT 14

T03714

5-epi-aristolochene synthase - common tobacco

N:Alternate names: sesquiterpene cyclase

C:Species: Nicotiana tabacum (common tobacco)

C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 29-Sep-1999

C:Accession: T03714

R:Facchini, P.J.; Chappell, J.

A:Title: Gene family for an elicitor-induced sesquiterpene cyclase in tobacco.

A:Reference number: Z15024; MUID:93066390; PMID:1438319

C:Accession: T03714

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-550 <FAC>

A:Cross-references: EMBL:L04680; NID:gl70342; PIDN:AAA19216.1; PID:g505588

A:Experimental source: strain NK326

C:Genetics:

A:Introns: 37/3; 126/1; 251/2; 324/2; 370/3; 452/3

C:Function:

A:Description: mediates the conversion of the isoprenoid intermediate farnesyl diphosphate to farnesyl pyrophosphate

C:Superfamily: vetispiradiene synthase 1

Query Match 15.9%; Score 633.5; DB 2; Length 550;

Best Local Similarity 29.0%; Pred. No. 1.6e-36;

Matches 153; Conservative 118; Mismatches 203; Indels 53; Gaps 15;

Qy 228 LFERLWAVDTVERLIGIDRYFKKEIKESLDYVYRWDAERGVGWARCNPDPVDDTAMGLR 287

Db 66 LADTLNIDIIERLIGISYHPEKEIDEILDQIYNQ-----NSNCN---DLCTSALQFR 114

Qy 288 ILRLHGVNVDVLENFDEKDFCFAGQTQIGVTDNLNLYRCSOVCPGKEKMEAAKT 347

Db 115 LLRQHGNISEIPEISKQDENGKPKESLASDVLGL---LNIYEAHVTHADDILEDALA 171

Qy 348 FTTNHLQNALAKNAFDMKWKOLPGEVEYAIKYPWHRSMRPLEARSYIEQFGSNDVWL 407

Db 172 FSTTHLESAAHP-----LKSPLREQVTHALEQCLHKGVPVETREFFI----- 213

Qy 408 GKTYYKMLYVSNKYLELAKLDNFNMVQALHOKETQIHVSWWRESGF-NDLFTTRQRPVEM 466

Db 214 -SSIYDKEQSKNNVLLRFKLDNFLLQMLHQLAEVSRWKKDLDFVTLTPYARDRVVEC 272

Qy 467 YFSVAVSMFEPEAFACRIAYAKTSCIAVLDDLYDTHGSLDDLKLFSEAVRRWDISVLDS 526

Db 273 YFWALGVYFEFOYQARVNLVKTISMSIVDDTDAYGTGVELEYATDAIGRWLINEIDR 332

Qy 527 VRDNOLKVCFLGLYNTVNGFGDGLKEQGRD-VLGYLKRWKWEGLLASYYTKEAASAKYV 585

Db 333 LPD-YMKISYKAILDLYKDYEKE-LSSAGRSHIVCHAIERKMEVVRNYNVESTWFIEGYM 390

Qy 586 PTFNEYENAKVSLATATVNLNSIFFTGEL-----LPDYLQOVDLRSFELHLVSLTGRLI 641

Db 391 PPVSEYLSN-----ALATTTYYLATTSYLGKMSATEQDFEWSKNPKILEASVILICRVI 445

Qy 642 NDTTYQAEARNRGELVSVQCYRENPECTEERALSHPYGIIDNALKELNWLANPASNA 701

Db 446 DDTATYEVKSGGGIATGIECYMEYDYGIST-KEAWAKFQNAETAWKNDINEGLLRP--- 501

Qy 702 PLCVRL--LPNTARVMOLFMYR-DGFGISDKEMKHVSRTLPDPV 745

Db 502 PVSTFELTPILNARIVEVTVIHNLGDGTHPEKVLKPHIINLLVDSI 548

## RESULT 15

T08174

sesquiterpene cyclase (EC 2.5.1.1) - pepper

C:Species: Capsicum annuum (pepper)

C&gt;Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: T08174

R:Back, K.; Shin, D.H.; He, S.

Plant Cell Physiol. 39, 899-904, 1998

A:Title: Cloning and bacterial expression of sesquiterpene cyclase, a key branch point enzyme in the biosynthesis of capsaicinoids in pepper

A:Reference number: Z16395; MUID:99033462; PMID:9816674

A:Accession: T08174

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-559 &lt;BAC&gt;

A:Cross-references: EMBL:AF061285; NID:g3108342; PIDN:AAC61260.1; PID:g3108343

C:Function:

A:Pathway: the synthesis of phytoalexin capsidiol

A:Note: induced by UV

C:Superfamily: vetispiradiene synthase 1

C:Keywords: isoprenoid biosynthesis; transferase

Query Match 15.9%; Score 633.5; DB 2; Length 559;

Best Local Similarity 30.7%; Pred. No. 1.6e-36;

Matches 163; Conservative 107; Mismatches 202; Indels 59; Gaps 15;

Search completed: August 17, 2004, 21:03:08

Job time : 16.7449 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:56:14 ; Search time 9.21557 Seconds  
(without alignments)  
4215.079 Million cell updates/sec

Title: US-10-041-007-39

Perfect score: 3972

Sequence: 1 MFQSMGDETNPSAYDTAWV.....GISDKMKDHVSRTLEDPVA 746

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1901	47.9	862	1	TASY_TAXBR Q41594 taxus brevi
2	1887	47.5	862	1	TASY_TAXHA Q93va3 taxus bacca
3	1858	46.8	862	1	TASY_TAXCH Q9f237 taxus chine
4	1194.5	30.1	637	1	TSD3 ABIGR Q22340 abies grand
5	1167.5	29.4	628	1	TSD2 ABIGR Q24474 abies grand
6	1151.5	29.0	627	1	TSD1 ABIGR Q24475 abies grand
7	647.5	16.3	554	1	DCS1 GOSAR Q39761 gossypium a
8	643.5	16.2	554	1	DCS2 GOSAR Q39760 gossypium a
9	643.5	16.2	554	1	DCS4 GOSAR Q49853 gossypium a
10	636.5	16.0	548	1	5BAS TOBAC Q40577 nicotiana t
11	631.5	15.9	555	1	DCS3 GOSAR Q43714 gossypium h
12	622.5	15.7	554	1	DCS1 GOSHI P93665 gossypium h
13	594.5	15.0	601	1	CASS_RICCO P59287 ricinus com
14	219.5	5.5	516	1	YAKT_RHISN P55538 rhizobium s
15	204.5	5.1	516	1	YI49 BRAJA Q45221 bradyrhizob
16	126.5	3.2	908	1	CNA3 MOUSE Q80xk6 mus musculu
17	122.5	3.1	757	1	YVDK BACSU Q06993 bacillus su
18	121	3.0	722	1	CNA3 HUMAN Q96by7 homo sapien
19	118.5	3.0	955	1	SVY BUCAI P57447 buchnera ap
20	118.5	3.0	1064	1	YV08 METJA Q60307 methanococ
21	117.5	3.0	775	1	REP_BHP1 P51711 bacteriopia
22	117	2.9	562	1	SVR_THETN Q8r786 thermoaner
23	116.5	2.9	1063	1	DPOM CLAPU P22373 claviceps p
24	115.5	2.9	1634	1	DPOL METJA Q58295 methanococ
25	114.5	2.9	1169	1	SMC METJA Q59037 methanococ
26	113.5	2.9	440	1	FULO_ARATH Q9sjp6 arabidopsis
27	113.5	2.9	799	1	LON_CAUCR P52977 caulobacter
28	113.5	2.9	7094	1	RIAB_CVBEN Q91a29 b replicase
29	113.5	2.9	7094	1	RIAB_CVBLU Q8v439 b replicase
30	111.5	2.8	748	1	Y875 METJA Q58285 methanococ
31	111.5	2.8	842	1	VJ27 YEAST P47050 saccharomyc
32	110.5	2.8	861	1	SVL_BUCBP P59433 buchnera ap
33	110.5	2.8	910	1	HUL5_YEAST P53119 saccharomyc

34	110.5	2.8	2215	1	SORL_MOUSE	O88307 m sortilin-
35	110.5	2.8	8797	1	SNEI_HUMAN	Qenf91 homo sapien
36	109.5	2.8	981	1	GLNE_HABIN	P44419 haemophilus
37	109.5	2.8	4568	1	DYHC CAEEL	Q1020 caenorhabdi
38	109	2.7	1020	1	CF6C HUMAN	O8nb25 homo sapien
39	109	2.7	1375	1	BNR1_YEAST	P40450 saccharomyc
40	108.5	2.7	1220	1	DXH8_HUMAN	Q14562 homo sapien
41	108.5	2.7	2283	1	DPOR_MOUSE	Q9wv77 mus musculu
42	108.5	2.7	7094	1	RIAB_CVBEM	O66198 b replicase
43	108	2.7	1111	1	EX5C BUCBP	O89ab4 buchnera ap
44	107.5	2.7	803	1	LON_HABIN	P43864 haemophilus
45	107	2.7	886	1	SYA_BARBA	P70865 bartonella

## ALIGNMENTS

RESULT 1						
TASY_TAXBR						
ID	TASY_TAXBR	STANDARD;	PRT;	862	AA.	
AC	Q41594; Q94FV8;					
DT	28-FEB-2003 (Rel. 41, Created)					
DT	28-FEB-2003 (Rel. 41, Last sequence update)					
DT	28-FEB-2003 (Rel. 41, Last annotation update)					
DB	Taxadiene synthase (SC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).					
GN	TDC1.					
OS	Taxus brevifolia (Pacific yew).					
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
OC	Spermatophyta; Coniferopsida; Coniferales; Taxaceae; Taxus.					
OX	NCBI_TaxID=46220;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=96199163; PubMed=8621577;					
RA	Wildung M.R., Croteau R.B.;					
RT	"A cDNA clone for taxadiene synthase, the diterpene cyclase that					
RT	catalyzes the committed step of taxol biosynthesis.";					
RL	J. Biol. Chem. 271:9201-9204(1996).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=21297238; PubMed=11404343;					
RA	Trapp S.C., Croteau R.B.;					
RT	"Genomic organization of plant terpene synthases and molecular					
RT	evolutionary implications.";					
RL	Genetics 158:811-832(2001).					
CC	-!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid					
CC	intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the					
CC	parent olefin with a taxane skeleton.					
CC	-!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene +					
CC	diphosphate.					
CC	-!- PATHWAY: Taxol biosynthesis; first step.					
CC	-!- SIMILARITY: Belongs to the terpene synthase family.					
CC						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.ebi.ac.uk/announcements/					
CC	or send an email to license@ebi.ac.uk).					
CC						
CC	EMBL; U48796; AAC49310.1; --					
DR	EMBL; AF26519; AAK83566.1; --					
DR	HSSP; Q40577; 5EAA.					
DR	InterPro; IPR008930; Terp_cyc toroid.					
DR	InterPro; IPR001906; Terp_synth-like.					
DR	InterPro; IPR005630; Terpene synth C.					
DR	InterPro; IPR008949; Terpenoid synth.					
DR	Pfam; PF01397; Terpene synth; 1.					
DR	Pfam; PF03936; Terpene synth C; 1.					
KW	Taxol biosynthesis; Lyase.					
FT	ACT_SITE 758					
FT	ACT_SITE 835					
FT	ACT_SITE 835 839					
FT	BY SIMILARITY.					
FT	BY SIMILARITY.					

FT	CONFLICT	148	148	L -> V (IN REF. 2).	
FT	CONFLICT	767	767	A -> V (IN REF. 2).	
SEQ	SEQUENCE	862 AA;	98303 MW;	9141B59780CD79A1 CRC64;	
Query Match					
Best Local Similarity 47.5%; Score 1901; DB 1; Length 862;					
Matches 358; Conservative 149; Mismatches 219; Indels 28; Gaps 9;					
QY	1	MFQSGDGETNPSAYDTAWARIPSI--DGSGAPQPPQTLQWILNNQLPDGSGWGEICIFL	58		
DB	126	MFNALGDDGISAYSAYDTAWARLATISDGSSEKPRFPQALNNVFNQLQDGSWGESHS	185		
QY	59	AYDRVLNTLACLLTLKWNKGDIOQKGVFVRKMEEMKDEADNRHPSGFEVPPAMLD	118		
DB	186	LCDRLLNTNSVIALSVKTHGSHQVQGAFFAENLRLLNEE-DELSPP-FEIIFFALLQ	243		
QY	119	EAKSGLDLPYHLPFISQIHQRKQKL-----QKIPNLVNHQTLALYSLEGLQDVV	171		
DB	244	KAKALGINLPYDLPFIKSLTTREARLTDVSAADNIPANMLN-----ALEGLEEVI	295		
QY	172	DWOEITNLSQSDGSFLSSPASTACVPMHTONKRCFLHFLNVLKSGDYVPCHYPLDLPER	231		
DB	296	DNKIMRFQSGDGSFLSSPASTACVPMHTONKRCFLHFLNVLKSGDYVPCHYPLDLPER	355		
QY	232	LWAVDTVERLGRDYPFKEIKESLDVYRYWDAERGVGWACNPIPDVDDTAMGLRLRL	291		
DB	356	LSLVNIEHLGIGRHFQKQKALDYVYRHW-SERGIGWGRDLSVPLDNTTALGLRLRT	414		
QY	292	HGYNVSSDVLNFRDEKGFPCFAGQTOIGVTNMLNLYRCSQVCPGKEKIMEAKTFTTN	351		
DB	415	HGYNVSSDVLNFRDEKGFPCFAGQTOIGVTNMLNLYRCSQVCPGKEKIMEAKTFTTN	474		
QY	352	HLQNALAKNAPDKWAKKDLPGVEYAIKYPHRSMPRLAARSYIQQFQSDNVWLKTV	411		
DB	475	YLREALA-----TKISTNTKLYKEIYVVEYPMHMSIPRLAARSYIQQFQSDNVWLKTV	529		
QY	412	YKMLVYSNEKYLELAKLDFNMVQALHQKQTHIVSWWRSGFNDLTFTTRQRPVEMVFA	471		
DB	530	YRMPSLNSKLELAKLDFNIVQSLHQELKLLTRWKGSGMADINFTHRVAEYVSSA	589		
QY	472	VSMFPEFAACRIAYAKTSCLAVILDDLYTHGSLDLDLKFSEAVRRWDISVLDSVRDQ	531		
DB	590	T--PEPEYSATRIAFKIGCLOQLVFDMDADIFATLDELKSFTEGVKRWDTSLIHEIPE-C	646		
QY	532	LKVCFLGLYNTVNGKQKGLKQGRDVLGYLRKWEGLLASITKEAWSAKVPFTNEY	591		
DB	647	MOTCFKWFKLMEEVNDVVKVQGRDMLAHIRKPMELYNFCYQBERWLEAGYIPTFEY	706		
QY	592	VENAKVSIATATVNLNIPFTGELLPDYILOQVDLRSKFLHLVSLTGLRLNDTKTYQAR	651		
DB	707	LKTYAISVGLGCTLOPILLMGLVKDDVVEKHYVPSNNFELVLSLWRLTNDTKTYQAK	766		
QY	652	NRGELVSSVQCYMRNPECTEALSHVVGIIIDNALKELNWLANPASHAPICVRRLLFN	711		
DB	767	ARGOQASGIACYMKNPGATEADAIGHICRVVDRLKEASFVFKPSPNDIPMGCKSFEN	826		
QY	712	TARVMQLFVMDYDGGISDKEMKOHVSRTLPDPV	745		
DB	827	LRLCVOIFPKYFDGYGIANEERIKDYIRKYVIDPI	860		
RESULT 2					
ID	TASY TAXBA	STANDARD;	PRT;	862 AA.	
AC	Q93YA3;				
DT	28-FEB-2003 (Rel. 41, Created)				
DT	28-FEB-2003 (Rel. 41, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Taxadiene synthase (EC 4.2.3.17) (Taxa-4(5),11(12)-diene synthase).				
GN	TDC1 OR TASF.				
OS	Taxus baccata (English yew).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Coniferales; Taxaceae; Taxus.				

OX	NCBI_TaxID=25629;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Goerhardt B.;				
RL	Thesis (2001), Technische Universitaet Berlin, Germany.				
CC	-!- FUNCTION: Catalyzes the cyclization of the ubiquitous isoprenoid intermediate geranylgeranyl diphosphate to taxa-4,11-diene, the parent olefin with a taxane skeleton.				
CC	-!- CATALYTIC ACTIVITY: Geranylgeranyl diphosphate = taxa-4,11-diene + diphosphate.				
CC	-!- PATHWAY: Taxol biosynthesis; first step.				
CC	-!- SIMILARITY: Belongs to the terpene synthase family.				
CC	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.ebi.ac.uk/announcements">http://www.ebi.ac.uk/announcements</a> or send an email to <a href="mailto:license@ebi.ac.uk">license@ebi.ac.uk</a> ).				
CC	EMBL; AJ320538; CAC42773.1; --				
DR	InterPro; IPR008930; Terp_cyc_toroid.				
DR	InterPro; IPR001906; Terp_synth-like.				
DR	InterPro; IPR005630; Terpene synth C.				
DR	InterPro; IPR008949; Terpenoid synth.				
DR	Pfam; PF01397; Terpene synth; 1.				
DR	Pfam; PF03936; Terpene synth; 1.				
KW	Taxol biosynthesis; Lyase.				
FT	ACT_SITE 758 758 BY SIMILARITY.				
FT	ACT_SITE 835 835 BY SIMILARITY.				
FT	ACT_SITE 839 839 BY SIMILARITY.				
SEQ	SEQUENCE 862 AA; 98048 MW; B37835AFFDC1D8FE CRC64;				
Query Match					
Best Local Similarity 47.1%; Pred. No. 7.7e-124;					
Matches 355; Conservative 152; Mismatches 219; Indels 28; Gaps 9;					
QY	1	MFQSGDGETNPSAYDTAWARIPSI--DGSGAPQPPQTLQWILNNQLPDGSGWGEICIFL	58		
DB	126	MFNALGDDGISAYSAYDTAWARLATISDGSSEKPRFPQALNNVFNQLQDGSWGESHS	185		
QY	59	AYDRVLNTLACLLTLKWNKGDIOQKGVFVRKMEEMKDEADNRHPSGFEVPPAMLD	118		
DB	186	LCDRLLNTNSVIALSVKTHGSHQVQGAFFAENLRLLNEE-DELSPP-FEIIFFALLQ	243		
QY	119	EAKSGLDLPYHLPFISQIHQRKQKL-----QKIPNLVNHQTLALYSLEGLQDVV	171		
DB	244	KAKALGINLPYDLPFIKSLTTREARLTDVSAADNIPANMLN-----ALEGLEEVI	295		
QY	172	DWOEITNLSQSDGSFLSSPASTACVPMHTONKRCFLHFLNVLKSGDYVPCHYPLDLPER	231		
DB	296	DNKIMRFQSGDGSFLSSPASTACVPMHTONKRCFLHFLNVLKSGDYVPCHYPLDLPER	355		
QY	232	LWAVDTVERLGRDYPFKEIKESLDVYRYWDAERGVGWACNPIPDVDDTAMGLRLRL	291		
DB	356	LSLVNIEHLGIGRHFQKQKALDYVYRHW-SERGIGWGRDLSVPLDNTTALGLRLRT	414		
QY	292	HGYNVSSDVLNFRDEKGFPCFAGQTOIGVTNMLNLYRCSQVCPGKEKIMEAKTFTTN	351		
DB	415	HGYNVSSDVLNFRDEKGFPCFAGQTOIGVTNMLNLYRCSQVCPGKEKIMEAKTFTTN	474		
QY	352	HLQNALAKNAPDKWAKKDLPGVEYAIKYPHRSMPRLAARSYIQQFQSDNVWLKTV	411		
DB	475	YLREALA-----TKISTNTKLYKEIYVVEYPMHMSIPRLAARSYIQQFQSDNVWLKTV	529		
QY	412	YKMLVYSNEKYLELAKLDFNMVQALHQKQTHIVSWWRSGFNDLTFTTRQRPVEMVFA	471		
DB	530	YRMPSLNSKLELAKLDFNIVQSLHQELKLLTRWKGSGMADINFTHRVAEYVSSA	589		
QY	472	VSMFPEFAACRIAYAKTSCLAVILDDLYTHGSLDLDLKFSEAVRRWDISVLDSVRDQ	531		
DB	590	T--PEPEYSATRIAFKIGCLOQLVFDMDADIFATLDELKSFTEGVKRWDTSLIHEIPE-C	646		
QY	532	LKVCFLGLYNTVNGKQKGLKQGRDVLGYLRKWEGLLASITKEAWSAKVPFTNEY	591		
DB	647	MOTCFKWFKLMEEVNDVVKVQGRDMLAHIRKPMELYNFCYQBERWLEAGYIPTFEY	706		
QY	592	VENAKVSIATATVNLNIPFTGELLPDYILOQVDLRSKFLHLVSLTGLRLNDTKTYQAR	651		
DB	707	LKTYAISVGLGCTLOPILLMGLVKDDVVEKHYVPSNNFELVLSLWRLTNDTKTYQAK	766		
QY	652	NRGELVSSVQCYMRNPECTEALSHVVGIIIDNALKELNWLANPASHAPICVRRLLFN	711		
DB	767	ARGOQASGIACYMKNPGATEADAIGHICRVVDRLKEASFVFKPSPNDIPMGCKSFEN	826		
QY	712	TARVMQLFVMDYDGGISDKEMKOHVSRTLPDPV	745		
DB	827	LRLCVOIFPKYFDGYGIANEERIKDYIRKYVIDPI	860		





CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in  
 CC response to insect attack or other injury. Involved in monoterpene  
 CC (C10) olefins biosynthesis.  
 CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate = limonene + diphosphate.  
 CC -!- COFACTOR: Manganese and potassium.  
 CC -!- PATHWAY: Oleoresinosis.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast.  
 CC -!- INDUCTION: By wounding.  
 CC -!- SIMILARITY: Belongs to the terpene synthase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF006193; AAB70907.1; -;  
 CC InterPro: IPR008930; Terp\_cyc toroid.  
 CC InterPro: IPR001906; Terp\_synth-like.  
 CC InterPro: IPR005630; Terpene synth C.  
 CC InterPro: IPR008949; Terpenoid synth.  
 CC Pfam: PF01397; Terpene synth\_1.  
 CC Pfam: PF03936; Terpene synth\_C; 1.  
 CC Lyase; Manganese; Transit peptide; Chloroplast.  
 CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).  
 CC CHAIN 1 637 BY SIMILARITY.  
 CC ACT\_SITE 533 610 BY SIMILARITY.  
 CC ACT\_SITE 610 610 BY SIMILARITY.  
 CC ACT\_SITE 614 614 BY SIMILARITY.  
 CC SEQUENCE 637 AA; 73477 MW; 8E80C9DDE886898 CRC64;  
 CC  
 CC Query Match 30.1%; Score 1194.5; DB 1; Length 637;  
 CC Best Local Similarity 41.9%; Pred. No. 1e-75;  
 CC Matches 244; Conservative 117; Mismatches 187; Indels 35; Gaps 9;  
 CC  
 CC 173 WQE--ITHLOS-RDGSFLSPASTAC-----VPMTONKRCILFLNVLKSGDVPCHY 224  
 CC 80 WEDDFIQSLSPYGGSSYSERAEVTVVEKMFNSIPNNREL-----FGS----- 124  
 CC  
 CC 225 PLDIFERLWAVDTVERLIGDRYFKKEIKESLDYVYVYWDAGRGVWACNPDPVDDTAM 284  
 CC 125 QNDLLTRLWVDSIERLIGDRHFQNEIRVALDYVSYWKEKGGICGRDSTFPDLNSTAL 184  
 CC 285 GLRLRLHGVNVSVDLENFRDEKGFCAFAGTQIGVTDN-LNYRCSQVCFPGKIME 343  
 CC 185 ALRTLRLHGVNVSVDLEVFYKDEKGFACPAITTEGQITRSVNLRYASLVAFPGKIME 244  
 CC 344 EAKPTTHNLQNALAKNNAFDKAVKCOLPGEVEYAIKYPWHRSPRLEARSYIEQFGSN 403  
 CC 245 EAEFTSASYLKKVLOK-----IPVSNLSGSEIYVLEYGHTNLPRLRARNYIEVYEOS 297  
 CC  
 CC 404 DWLGLKTVKMLYVSEKYLELAKLDFNNVALHOKETQIHVSWWRSEFNDLTETTRQP 463  
 CC 298 GY---ESLNEMPYMMKLLQLAKLEFNIFSLQRLRELQISRWKESGSSQLTTRHRH 354  
 CC 464 VEMYSVAVSMPEPFAACRIAYAKTSCLAVILDDLDTYTHGSLDLKLPSEAVRWIDSV 523  
 CC 355 VEYITMASICISMLPKHSAFRMEFVCHLVTLVLDIDYDTFGTMNELQLFTDAIKRWDLST 414  
 CC 524 LDSVRDNLKVCFLGTYNTVNGFGDKLKEQGRDVLGYLRKVEGLLASYTKEAESAAK 583  
 CC 415 TRWLPE-YMKGYMOLYOCINEMVEAEKTOGRDMLNYIQNAWEALFTFMQEAKEWISS 473  
 CC 584 YVPTNEVYENAKVSTALATVVLNSITFFTGELLPDYILQVDDLSKFLHLVSLTGRIND 643  
 CC 474 VLPTEEVILKNAKVGSGRIATLPILTLDPVLPDYILQEIYDPSFNFELASILLRGD 533  
 CC 644 TKTYQAERNRGLVSVQCYMRENPCTEELASHVYGIIDNALKELNLANPASNAPL 703  
 CC 534 TRCYKADRARGEASAIKCYMKDHPGISEBDALNHNAMISDAIRELNWELLRPSKSPI 593

QY 704 CVRRLEFNTARVQMFYMYRDGFGISDKEMKDHVSRITLDPVA 746  
 Db 594 SSKKHAFTITRAFHVYKYRDGYTVSNNETKNLVKMTVLEPLA 636  
 RESULT 5  
 TSD2 ABIGR STANDARD; PRT; 627 AA.  
 AC 024474;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Myrcene synthase, chloroplast precursor (EC 4.2.3.15).  
 GN AG2.2.  
 OS Abies grandis (Grand fir).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 OX NCBI\_TaxID=46611;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=97413772; PubMed=9268308;  
 RA Bohmann J., Steele C.L., Croteau R.;  
 RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation,  
 RT characterization, and functional expression of myrcene synthase, (-)-  
 RT (4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";  
 RL J. Biol. Chem. 272:21784-21792(1997).  
 CC -!- FUNCTION: Involved in defensive oleoresin formation in conifers in  
 CC response to insect attack or other injury. Involved in monoterpene  
 CC (C10) olefins biosynthesis.  
 CC -!- CATALYTIC ACTIVITY: Geranyl diphosphate = myrcene + diphosphate.  
 CC -!- COFACTOR: Manganese and potassium.  
 CC -!- PATHWAY: Oleoresinosis.  
 CC -!- SUBCELLULAR LOCATION: Chloroplast.  
 CC -!- INDUCTION: By wounding.  
 CC -!- SIMILARITY: Belongs to the terpene synthase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U87908; AAB71084.1; -;  
 CC HSP; Q40577; SEAS.  
 CC InterPro: IPR008930; Terp\_cyc toroid.  
 CC InterPro: IPR001906; Terp\_synth-like.  
 CC InterPro: IPR005630; Terpene synth C.  
 CC InterPro: IPR008949; Terpenoid synth.  
 CC Pfam: PF01397; Terpene synth\_1.  
 CC Pfam: PF03936; Terpene synth\_C; 1.  
 CC Lyase; Manganese; Transit peptide; Chloroplast.  
 CC TRANSIT 1 ? CHLOROPLAST (POTENTIAL).  
 CC CHAIN 1 627 MYRCENE SYNTHASE.  
 CC ACT\_SITE 523 523 BY SIMILARITY.  
 CC ACT\_SITE 600 600 BY SIMILARITY.  
 CC ACT\_SITE 604 604 BY SIMILARITY.  
 CC SEQUENCE 627 AA; 72478 MW; 2E0DA492E0C971FD CRC64;  
 CC  
 CC Query Match 29.4%; Score 1167.5; DB 1; Length 627;  
 CC Best Local Similarity 43.2%; Pred. No. 7.6e-74;  
 CC Matches 224; Conservative 104; Mismatches 173; Indels 17; Gaps 5;  
 CC  
 CC 227 DLFERLWAVDTVERLIGDRYFKKEIKESLDYVYVYWDAGRGVWACNPDPVDDTAMGL 286  
 CC 122 DLQRLWIVDSVERLIGARHFKNEITSALDYVRYWE-ENGICGRDSTVTDLNLSTALGF 180  
 CC 287 RLRLHGVNVSVDLENFRDEKGFCAFAGTQIGVTDNLYRCSQVCFPGKIMEAK 346  
 CC 181 RTLRHLGYTVSPVLKAFQDQNGQFVCSQGTGERSVNLRYASLVAFPGKIMEAE 240  
 CC 347 TPTTNHLQNALAKNNAFDKAVKCOLPGEVEYAIKYPWHRSPRLEARSYIEQFGSN-DV 405



Db 241 IFSTRYKLEALQK-----IPVSAISOEIRFVMEYGHWTNLPRLRARNYIDTLEKOTSA 293  
Qy 406 WLGTVTYKMLYVSNKYLELAKLDFNMVQALHOKETQHIVSWWRESGFNDLTFTRQREVE 465  
Db 294 WLKNK-----ACKKLELAKLFNFNSLQKELQVLLRWKESDLPKLTFAHRHVE 346  
Qy 466 MYFSVAVSMPEPEFAACRIAYAKTSCLAVIDDLIDYTHGSLDDLLKLPSEAVRRWDISVLD 525  
Db 347 FYTLASCIADIPKHSAPFLGPAKCHLVTVLDDIYDTFTGTDLELELFTSAIKRWNSSBIE 406  
Qy 526 SVRDNLKVCFLGLYNTVNGFGDKLQEGRDVLGRLKRWEGLLASVTKAEWSAAKV 585  
Db 407 HLPB-YMKCVVMVFETVNEUTRAEKTQGRNTLVYRKAWAEVDSYWEAKWISNGYL 465  
Qy 586 PTFNEYVENAKVSIATATVVLNSIFFTGELLPDYILOQVLDLRSKFLHLVSLTGRLLINDTK 545  
Db 466 PMPEYHENGKVSAYRVATLQPLTLNWLDPVILKIDGIDPFSPFNDLASSFLRLRGDTR 525  
Qy 646 TYQABNRGELVSSVOCYMRNPECTEBEALSHVYGIIDNALKELNWLNPASNAPLCV 705  
Db 526 CYKADRDGEASCISSCYMKDNPSTEDALNHINAMVNDIIEKLNWELLRSNDNIPMLA 585  
Qy 706 RLLFNTARVMOLFMYRDRDGFISDKEMKHVSRTLPD 743  
Db 586 KGHAFDITRALHLYIYRDGFSVANKETKULVMTILLE 623

## RESULT 6

TS01\_ABIGR STANDARD; PRT; 628 AA.  
AC Q24475;  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pinene synthase, chloroplast precursor (EC 4.2.3.14) (Beta-geraniolene synthase) ((-)-(1S,5S)-pinene synthase).  
DE AG3.18.  
GN Abies grandis (Grand fir).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
OX NCBI\_TaxID=46611;  
RN [1]  
SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RX MEDLINE=97413772; PubMed=9268308;  
RA Bohlmann J., Steele C.L., Croteau R.;  
RT "Monoterpene synthases from grand fir (Abies grandis). cDNA isolation, characterization, and functional expression of myrcene synthase, (-)-(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase.";  
RT J. Biol. Chem. 272:21784-21792(1997).  
RL  
CC -1- FUNCTION: Involved in defensive oleoresin formation in conifers in response to insect attack or other injury. Involved in monoterpene (C10) olefins biosynthesis. A mixture of alpha- and beta-pinene is produced by this enzyme.  
CC -1- CATALYTIC ACTIVITY: Geranyl diphosphate = pinene + diphosphate.  
CC -1- COFACTOR: Manganese and potassium.  
CC -1- PATHWAY: Oleoresinosis.  
CC -1- SUBCELLULAR LOCATION: Chloroplast.  
CC -1- INDUCTION: By wounding.  
CC -1- SIMILARITY: Belongs to the terpene synthase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U87909; AAB71085.1; --  
CC HSPSP; Q40577; SEAT.  
DR InterPro; IPR008930; Terp\_cyc\_toroid.  
DR InterPro; IPR001906; and characterization of (+)-delta-cadinene

DR InterPro; IPR005630; Terpene synth C.  
DR InterPro; IPR008949; Terpenoid synth.  
DR Pfam; PF01397; Terpene synth; 1.  
KW Pfam; PF03936; Terpene synth C; 1.  
KW Lyase; Manganese; Transit peptide; Chloroplast.  
FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).  
FT CHAIN ? 628 PINENE SYNTHASE.  
FT ACT\_SITE 524 524 BY SIMILARITY.  
FT ACT\_SITE 601 601 BY SIMILARITY.  
FT ACT\_SITE 605 605 BY SIMILARITY.  
SQ SEQUENCE 628 AA; 71505 MW; 23DBB788F3C8072C CRC64;  
Query Match 29.0%; Score 1151.5; DB 1; Length 628;  
Best Local Similarity 44.0%; Pred. No. 1e-72;  
Matches 230; Conservative 97; Mismatches 179; Indels 17; Gaps 6;  
Qy 225 PL-DLFEFLWAVDTVERLGIDRYFKKEIKESLDYVYVYWDAGRGVWACNPIPDVDDTA 283  
Db 119 PLNDLIQRLWIVDSLERLGIHRHFDEIKSALDYVYVYVW-GENGICGRESVVTDLNSTA 177  
Qy 284 MGLRLRLHGVNSVDVLENFRDEKDPFCFAG-QTQIGVTNMLNLYRSCVCPGPKIM 342  
Db 178 LGLRLRLHGVNSVDVLENFRDEKDPFCFAG-QTQIGVTNMLNLYRSCVCPGPKIM 237  
Qy 343 REAKTFTTNHLQNALAKNNAFDKMAVKKDLPGVEVYAIKYPWHRSMRLEARSVIEQFGS 402  
Db 238 DEARIFSTKYLKEALQK-----IPVSSLSREIGDVLSEYGHWTNLPRLRARNYIQVFGQ 290  
Qy 403 NDVWLKGTVYKMLYVSNKYLELAKLDFNMVQALHOKETQHIVSWWRESGFNDLTFTRQ 462  
Db 291 D-----TENTKSYVKSCKLLELAKLFNIFQSLQKRELSLVRWKESGFPEMTCRHR 344  
Qy 463 PVEMYFSVAVSMPEPEFAACRIAYAKTSCLAVIDDLIDYTHGSLDDLLKLPSEAVRRWDIS 522  
Db 345 HVEYTYTLASCIADIPKHSAPFLGPAKCHLVTVLDDIYDTFTGTDLELELFTATMKRWDP 404  
Qy 523 VLDSDVRDNLKVCFLGLYNTVNGFGDKLQEGRDVLGRLKRWEGLLASVTKAEWSAA 582  
Db 405 SIDCLPE-YMKGVYIAVDVTNEMAREAEAGQRTLTAREAWAEYVDSYMEARMIAT 463  
Qy 583 KVPFTFNEVENAKVSIATATVVLNSIFFTGELLPDYILOQVLDLRSKFLHLVSLTGRLLN 642  
Db 464 GVLPSFDEYENGKVSCHGRISALQPLTMDIPFPDHLKEVDPPSKLNDLACAILRLG 523  
Qy 643 PTKTYOABNRGELVSSVOCYMRNPECTEBEALSHVYGIIDNALKELNWLNPASNA 702  
Db 524 DTRCYKADRDGEASCISSCYMKDNPSTEDALNHINAMVNDIIEKLNWELLKPKDINVP 583  
Qy 703 LCVRRLLFNTARVMOLFMYRDRDGFISDKEMKHVSRTLPD 745  
Db 584 ISAKKHAFDITRALHLYIYRDGFSVANKETKULVMTILLE 626  
RESULT 7  
DCS1\_GOSAR STANDARD; PRT; 554 AA.  
AC Q39761;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE (+)-delta-cadinene synthase isozyme Xc1 (EC 4.2.3.13) (D-cadinene synthase).  
OS Gossypium arboreum (Tree cotton).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
OX NCBI\_TaxID=29729;  
RN [1]  
SEQUENCE FROM N.A.  
RC STRAIN=cv. Nanking;  
RX MEDLINE=96132653; PubMed=8554317;  
RA Chen X.-Y., Chen Y., Heinsteins P., Davisson V.J.;  
RT "Cloning, expression, and characterization of (+)-delta-cadinene

RT synthase: a catalyst for cotton phytoalexin biosynthesis.";

RL Arch. Biochem. Biophys. 324:255-266(1995).

CC -I- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl

CC diphosphate (fpp) to (+)-delta cadinene.

CC -I- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-

CC delta-cadinene + diphosphate.

CC -I- PATHWAY: Phytoalexin gossypol and lacinilene C biosynthesis;

CC first (committed) step

CC -I- SIMILARITY: Belongs to the terpene synthase family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; U23206; AAA93064.1; -

DR PIR; S68365; S68365.

DR HSP; Q40577; SEAU.

DR InterPro; IPR008930; Terp\_cyc\_toroid.

DR InterPro; IPR001906; Terp\_synth-like.

DR InterPro; IPR005630; Terpene synth C.

DR InterPro; IPR008949; Terpenoid synth.

DR Pfam; PF01397; Terpene synth\_1.

DR Pfam; PF03936; Terpene synth\_C; 1.

KW Lyase; Multigene family.

FT ACT\_SITE 451 BY SIMILARITY.

FT ACT\_SITE 527 BY SIMILARITY.

FT ACT\_SITE 531 BY SIMILARITY.

FT ACT\_SITE 531 BY SIMILARITY.

SQ SEQUENCE 554 AA; 64137 MW; 5906922DEDF9DCAF CRC64;

Query Match 16.3%; Score 647.5; DB 1; Length 554;

Best Local Similarity 30.9%; Pred. No. 1.2e-37;

Matches 162; Conservative 103; Mismatches 210; Indels 49; Gaps 12;

QY 230 ERLAVDTVERIGIDRYFKKEIKESLDYVY-RYWDAGRGVWGCNPIPDVDDTAMGLRI 288

DB 72 QKLAFTDSVQRLLGVSYHFTKEIEDELENIYHNNDNAEN-----DLTYTSLRFL 120

QY 289 LRLHGVNVSVDVLEFRDEKGFPCFAGQTQIGVTDN-----LNLVRCQVCPFGKIMEE 344

DB 121 LREHGVNVSVDVFNKFDQGNF-----KSVTSDEVGLLELYQASYLRVHGEDI 173

QY 345 AKTFTTNHQLNALKNAPDKWAKKDLPGVEYAIKYPWHRSMRPLEARSYIEQFGSND 404

DB 174 AISFTTHLSLAVA-----SLDYPLESEVSHALKQSIIRGLPRVEARHYL----- 218

QY 405 VMLGKTVYKMLYVSNKYLELAKLDPMVQALHOKETQHVSWWRSGF-NDLTFTTRQP 463

DB 219 -----SVQDIESHNKALDEFKIDFNMLQFLLRKLSEICRWKKDLDFQKLPYARDRV 273

QY 464 VEMYSFVAVSMPEPFAACRIAYAKTSCILAVILDDLYDTHGSLDDIKLPSEAVRRWDISV 523

DB 274 VEGYFWISGVYEPQYSLGRKMLTKVIASIVDDYDSYATVEELIPTNAIERWDIKC 333

QY 524 LDSVRDNLKVCFLGLYNTVNGFGKGLKEQGRDV-LGYLRKWEGGLLASYTEKAEWSAA 582

DB 334 IDEIPE-YMKPSYKALLDYYEEM-VQLVAHGQRVQVEYAKNMTIRLAQSYLVEAKWTIQ 391

QY 583 KVPFTFNEVENAKYSIALATVVLNSIIFTGELLPPYIIQQVDLRSKFLHLVSLTGRLLN 642

DB 392 NTKPSFEPEKANALPTCGYAMLATISFVGMGDIIVTPTFKWAASDPKIIQASTIICRFMD 451

QY 643 DTKTYQAEENRGLSVSSVOCYRENPECTEEESALSHVYGIIDNALKELNWLNPASNPAP 702

DB 452 DVAEHKFKHRREDDCSAISCYMEY-GVTAQRAYDVFNKHVESAWKDLNQEFKLP-TEMP 509

QY 703 LCVRRLLFNARVMQLFYMYRGGFGISKEMKHDSRTLPDPVA 746

DB 510 TEVLNRLNLARVMVLYREGDGYTVYVGAAGKGGITSLIETPIA 553

RESULT 8

DCS2\_GOSAR STANDARD; PRT; 554 AA.

AC Q39760;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE (+)-delta-cadinene synthase isozyme Xc14 (EC 4.2.3.13) (D-cadinene synthase).

OS Gossypium arboreum (Tree cotton).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Malvales; Malvaceae; Gossypium.

OX NCBI\_TaxID:29729;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Nanking.

RX MEDLINE=96132653; PubMed=8554317;

RA Chen X.-Y., Chen Y., Heinsteins P., Davison V.J.;

RT "Cloning, expression, and characterization of (+)-delta-cadinene synthase: a catalyst for cotton phytoalexin biosynthesis.";

RL Arch. Biochem. Biophys. 324:255-266(1995).

CC -I- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl diphosphate (fpp) to (+)-delta cadinene.

CC -I- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-

CC delta-cadinene + diphosphate.

CC -I- PATHWAY: Phytoalexin gossypol and lacinilene C biosynthesis;

CC first (committed) step.

CC -I- SIMILARITY: Belongs to the terpene synthase family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; U23205; AAA93065.1; -

DR PIR; S68366; S68366.

DR HSP; Q40577; SEAU.

DR InterPro; IPR008930; Terp\_cyc\_toroid.

DR InterPro; IPR001906; Terp\_synth-like.

DR InterPro; IPR005630; Terpene synth C.

DR InterPro; IPR008949; Terpenoid synth.

DR Pfam; PF01397; Terpene synth\_1.

DR Pfam; PF03936; Terpene synth\_C; 1.

KW Lyase; Multigene family.

FT ACT\_SITE 451 BY SIMILARITY.

FT ACT\_SITE 527 BY SIMILARITY.

FT ACT\_SITE 531 BY SIMILARITY.

FT ACT\_SITE 531 BY SIMILARITY.

SQ SEQUENCE 554 AA; 64158 MW; A897466580F6B2B CRC64;

Query Match 16.2%; Score 643.5; DB 1; Length 554;

Best Local Similarity 30.8%; Pred. No. 2.3e-37;

Matches 161; Conservative 103; Mismatches 212; Indels 47; Gaps 10;

QY 230 ERLAVDTVERIGIDRYFKKEIKESLDYVY-RYWDAGRGVWGCNPIPDVDDTAMGLRI 288

DB 72 QKLAFTDSVQRLLGVSYHFTKEIEDELENIYHNNDNAEN-----DLTYTSLRFL 120

QY 289 LRLHGVNVSVDVLEFRDEKGFPCFAGQTQIGVTDN-----LNLVRCQVCPFGKIMEE 344

DB 121 LREHGVNVSVDVFNKFDQGNF-----KSVTSDEVGLLELYQASYLRVHGEDI 173

QY 345 AKTFTTNHQLNALKNAPDKWAKKDLPGVEYAIKYPWHRSMRPLEARSYIEQFGSND 404

DB 174 AISFTTHLSLAVA-----SLDYPLESEVSHALKQSIIRGLPRVEARHYL----- 218

QY 405 VMLGKTVYKMLYVSNKYLELAKLDPMVQALHOKETQHVSWWRSGF-NDLTFTTRQP 463

DB 219 -----SVQDIESHNKALDEFKIDFNMLQFLLRKLSEICRWKKDLDFQKLPYARDRV 273

QY 464 VEMFVSVAVMFEPFAACRIAYAKTSCLAVILDYDTHGSLDCLKLFSEAVRRWDISV 523  
 DB 274 VEGYFWISGVVFEQYSLGRKMLTKVIAMASIVDDYDVSATYEEELIPYTKAIERWDIKC 333  
 QY 524 LDSVRDNLKVCFLGLNTVNGFKDGLKEGDRDVLGVLKVKWGLLASYTKAEWSAAK 583  
 DB 334 IDELPE-YMKPSYKALLDVVEEMEQLVAKHGRQYRVEYAKNAMIRLAQSYLVEARWTLQ 392  
 QY 584 YVTFENYVENAKVSIATATVNLNSIFFTGELLDPYILQQVDLSKFLHLVSLTGRLLND 643  
 DB 393 YKPSFEEFKANALPTCGYAMLAITSFVGMGDIVTPTFKWAANDPKIIQASTIICRFMD 452  
 QY 644 TKTYQAEERNRGELVSSVQCYMRENPCTEEBALSHVYGIIDNALKELNWLANSAPNAPL 703  
 DB 453 VAHKPKHRRDDCSAIECYMEY-GVTAQAEYDVFNKGVESAKMDVKNKFLKP-TEMP 510  
 QY 704 CVRRLLFNTARVMOLFMYRDRGFGISDKEMKHVSRTLFDPVA 746  
 DB 511 EVLNRSNLARVMDVLYREGDGYTVGKAAGGITSLLIEBPA 553

## RESULT 9

DCS4\_GOSAR STANDARD; PRT; 554 AA.  
 AC O49853;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE (+)-delta-cadinene synthase isozyme C2 (EC 4.2.3.13) (D-cadinene synthase).  
 DE synthase).  
 GN CAD1-C2.  
 OS Gossypium arboreum (Tree cotton).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.  
 OX NCBI\_TaxID=29729;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nanking.  
 RA Meng Y., Jia J., Liu C., Liang W., Zhou X., Heinstein P., Chen X.-Y.;  
 RA Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RL  
 CC -I- FUNCTION: Responsible for the cyclization of trans,trans-farnesyl diphosphate (FPP) to (+)-delta-cadinene.  
 CC -I- CATALYTIC ACTIVITY: 2-trans,6-trans-farnesyl diphosphate = (+)-delta-cadinene + diphosphate.  
 CC -I- PATHWAY: Phytoalexins gossypol and lacinilene C biosynthesis;  
 CC first (committed) step.  
 CC -I- SIMILARITY: Belongs to the terpene synthase family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Y16432; CAA76223.1; -  
 CC HSP; Q40577; SEAU.  
 DR InterPro; IPR008930; Terp\_cyc\_toroid.  
 DR InterPro; IPR001906; Terp\_synth-like.  
 DR InterPro; IPR005630; Terpene synth C.  
 DR InterPro; IPR008949; Terpenoid synth.  
 DR Pfam; PF01397; Terpene synth; I.  
 DR Pfam; PF03936; Terpene synth; C; 1.  
 DR Lyase; Multigene family.  
 KW ACT\_SITE 451 451 BY SIMILARITY.  
 FT ACT\_SITE 527 527 BY SIMILARITY.  
 FT ACT\_SITE 531 531 BY SIMILARITY.  
 SQ SEQUENCE 554 AA; 64117 MW; 35DD66D3E838AAC CRC64;

Query Match 16.2%; Score 643.5; DB 1; Length 554;

Best Local Similarity 30.3%; Pred. No. 2.3e-37;  
 Matches 160; Conservative 105; Mismatches 206; Indels 57; Gaps 13;  
 QY 230 BRLWAVDVERLIGDRYFKKIKESLDYVY-RYMDAERGVGWARCNPIDVDDDTAMGLRI 288  
 DB 72 OKLAFIDSQVLGVSYHFTKEIDLENIYHNNDAAEN-----DLYTSLRPL 120  
 QY 289 LRLHGVNVSVDLENFRDEKGFPCFAGOTIGVTDN-----LNLRCOSQVCPFGKIMEE 344  
 DB 121 LREHGVNVSVDLENFRDEKGFPCFAGOTIGVTDN-----LNLRCOSQVCPFGKIMEE 344  
 QY 345 AKTPTTNHLQNALAKNAPDKWAVKDLPGVEYAIKYPWHRSMRPLREARSIEQFGSND 404  
 DB 174 AISFTTNHLSAVS-----SLDHPLESEYSHALKOSIRRGRLPRVEARHYL----- 218  
 QY 405 VMLGTVYKMLVYSNEKYLELAKLDFNMVQALHOKETOHIVSWWRSEGF-NDLTFTQRP 463  
 DB 219 -----SVYQDIESHNKALLEFAKIDFNMLQFLHRLKELSEICRWKWDLDLQKLPYARDV 273  
 QY 464 VEMFVSVAVMFEPFAACRIAYAKTSCLAVILDYDTHGSLDCLKLFSEAVRRWDISV 523  
 DB 274 VEGYFWISGVVFEQYSLGRKMLTKVIAMASIVDDYDVSATYEEELIPYTKAIERWDIKC 333  
 QY 524 LDSVRDNLKVCFLGLNTVNGFKDGLKEGDRDVLGVLKVKWGLLASYTKAEWSAAK 583  
 DB 334 IDELPE-YMKPSYKALLDVVEEMEQLVAKHGRQYRVEYAKNAMIRLAQSYLVEARWTLQ 392  
 QY 584 YVTFENYVENAKVSIATATVNLNSIFFTGELLDPYILQQVDLSKFLHLVSLTGRLLND 643  
 DB 393 YKPSFEEFKANALPTCGYAMLAITSFVGMGDIVTPTFKWAANDPKIIQASTIICRFMD 452  
 QY 644 TKTYQAEERNRGELVSSVQCYMRENPCTEEBALSHVYGIIDNALKELNWLANSAPNAPL 703  
 DB 453 VAHKPKHRRDDCSAIECYMEY-GVTAQAEYDVFNKGVESAKMDVKNKFLKP-TEMP 510  
 QY 704 CVRRLLFNTARVMOLFMYRDRGFGISDKEMKHVSRTLFDPVA 746  
 DB 511 EVLNRSNLARVMDVLYREGDGYTVGKAAGGITSLLIEBPA 553

## RESULT 10

SEAS\_TOBAC STANDARD; PRT; 548 AA.  
 AC Q40577;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Aristolochene synthase (EC 4.2.3.9) (5-epi-aristolochene synthase) (EAS).  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamids; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 56-73.  
 RC STRAIN=cv. NK326;  
 RX MEDLINE=93066390; PubMed=1438319;  
 RA Facchini P.J., Chappell J.;  
 RT "Gene family for an elicitor-induced sesquiterpene cyclase in tobacco."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:11088-11092 (1992).  
 RN [2]  
 RX MEDLINE=974442533; PubMed=9295271;  
 RA Starks C.M., Back K., Chappell J., Noel J.P.;  
 RT "Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase."  
 RL Science 277:1815-1820 (1997).  
 CC -I- FUNCTION: Catalyzes the cyclization of trans,trans-farnesyl diphosphate (FPP) to the bicyclic intermediate 5-epi-aristolochene, initial step in the conversion of FPP to the sesquiterpenoid antifungal phytoalexin capsidiol.

CC -!- CATALYTIC ACTIVITY: Trans,trans-farnesyl diphosphate =  
CC aristolochene + diphosphate.  
CC -!- COPACTOR: Binds 3 magnesium ions per subunit.  
CC -!- SUBUNIT: Monomer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- INDUCTION: By fungal elicitor.  
CC -!- SIMILARITY: Belongs to the terpene synthase family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL; L04680; AAA192116.1; -;  
DR PIR; T03714; T03714.  
DR PDB; SEAS; 15-OCT-97.  
DR PDB; SEAT; 12-NOV-97.  
DR PDB; SEAU; 08-APR-98.  
DR InterPro; IPR008930; Terp\_cyc\_toroid.  
DR InterPro; IPR001906; Terp\_synth-like.  
DR InterPro; IPR005630; Terpene synth C.  
DR InterPro; IPR008949; Terpenoid synth.  
DR Pfam; PF01397; Terpene synth; 1.  
DR Pfam; PF03936; Terpene synth; C; 1.  
KW Lyase; Magnesium; 3D-structure.  
FT ACT\_SITE 273 273 PROTON ACCEPTOR.  
FT ACT\_SITE 444 444  
FT ACT\_SITE 520 520  
FT ACT\_SITE 525 525  
FT CONFLICT 42 42 Y -> Y (IN REF. 1).  
FT CONFLICT 44 44 K -> Q (IN REF. 1).  
FT CONFLICT 55 55 N -> S (IN REF. 1).  
FT CONFLICT 62 62 M -> R (IN REF. 1).  
FT CONFLICT 73 73 T -> I (IN REF. 1).  
FT CONFLICT 89 89 D -> E (IN REF. 1).  
FT CONFLICT 388 388 T -> M (IN REF. 1).  
FT HELIX 26 29  
FT HELIX 36 57  
FT TURN 58 58  
FT TURN 60 61  
FT HELIX 64 76  
FT TURN 77 78  
FT HELIX 80 83  
FT HELIX 84 97  
FT TURN 104 116  
FT TURN 117 118  
FT STRAND 123 129  
FT TURN 130 130  
FT TURN 132 133  
FT STRAND 136 136  
FT HELIX 138 142  
FT HELIX 144 154  
FT TURN 155 156  
FT TURN 159 160  
FT HELIX 162 164  
FT TURN 165 166  
FT HELIX 167 178  
FT HELIX 179 181  
FT TURN 184 185  
FT HELIX 186 195  
FT TURN 199 200  
FT TURN 203 213  
FT TURN 214 214  
FT HELIX 215 217  
FT TURN 219 220  
FT HELIX 223 224  
FT TURN 225 225  
FT HELIX 256 259  
FT TURN 261 262  
FT HELIX 267 277

FT	HELIX	281	283
FT	HELIX	284	306
FT	TURN	307	307
FT	HELIX	310	322
FT	TURN	323	323
FT	HELIX	326	330
FT	HELIX	333	354
FT	TURN	355	357
FT	HELIX	359	361
FT	HELIX	362	385
FT	TURN	386	386
FT	HELIX	391	398
FT	TURN	399	399
FT	HELIX	400	402
FT	HELIX	404	413
FT	TURN	414	414
FT	TURN	416	417
FT	HELIX	420	427
FT	TURN	428	428
FT	HELIX	431	454
FT	TURN	455	456
FT	TURN	458	459
FT	HELIX	461	469
FT	TURN	470	470
FT	HELIX	473	494
FT	TURN	495	495
FT	HELIX	503	505
FT	TURN	506	506
FT	HELIX	507	519
FT	TURN	520	520
FT	HELIX	535	542
FT	TURN	543	543
SQ	SEQUENCE	548 AA;	62973 MW; 9PELC59CFIA68BF1 CRC64;

Query Match 16.0%; Score 636.5; DB 1; Length 548;  
Best Local Similarity 28.9%; Pred. No. 6.9e-37;  
Matches 153; Conservative 119; Mismatches 204; Indels 53; Gaps 15;

QY	226	LDLPERLWAVTVRLGIDRYFKKEIKESLDYVYRYWDAERGVGWARCNPDPVDDTAMG	285
DB	62	MKLADTLNLIDTIERLGISYHFEKEIDDLQLYNQ-----NSNCN---DLCTSLAQ	110
QY	286	LRILRLHGVNVDVLENERDEKDFCFAGOTQIGVTNLDNLYRCSQVCFPGKEIMEEA	345
DB	111	FRLLRQHGFNISPEIFSKFQDENGKFKESLASDVLGL---LNLYEASHVTHADDILEDA	167
QY	346	KTFNTNHLQNALAKONNAFDKVAVKDLPQVEYVAIKYPWHRSPRLREARSYIEQFGSNDV	405
DB	168	LAFSTTHLESAAPH-----LKSPLREQVTHALEQCLHKGVPVETREFFI-----	211
QY	406	WLGKTVYKMLYVSNKYLELAKLDFNMVQALHOKETQHVIVSMWRSGF-NDLFTTRQPV	464
DB	212	---SSYIDKEQSKNNVLLRFKDLDFNLQMLHKQELAQVSRWKKDLDFVTTLPYARDVV	268
QY	465	EMYFSVAVSMFEPEFAACRIAYAKTSCLAVILDLDYDTHGSLDLDLKLFSBARRWDISVL	524
DB	269	ECYFWALGVYFEPQYSQARVLMVKTISMSISVDDTFDAYGVTKLEBAYTDAIQKWDINEI	328
QY	525	DSVRDNLKVCFLGLYNTVNGFGKQGLKEQGRD-VLGYLKKWEGLLASYTKAEWSAAK	583
DB	329	DRLPD-YMKISYKAILDLKYDYEK-LSSAGRSHIVCHAIERKKEVVRNRYNVESTWFIEG	386
QY	584	YVPTENEYVENAKVSIATATVNLNLSIFFTGEL---LPDYILOQVDLRSKFLHVSUTGR	639
DB	387	YTPPVSEYLSN-----ALATTTYYLATTSLYLGKMSATEQDFEWSKNPKILASVILCR	441
QY	640	LINDTKTYQAEARNRGELSVSQCYMRENPECTEEALSHPYIGIDNALKEINMELANPAS	699
DB	442	VIDDITATYVEKSRGQIATGIECCMDYDGI-ST-KEAMAKFQNMATKAWKIDINEGLLRP--	498
QY	700	NAPLCVRRLL--LFTARVNLQFMYR-DGFGISDKEMKDHVSRTLPDPV	745
DB	499	-TPVSTEFLLPILNLARIVEVTYIHLNLDGYTHPEKVLKPHIINLNLVDSI	546





```
QY 637 TGRINDTKYQAEARNRGELVSSVOCYMRNPECTEEBALSHTVYGIINALKELNWEIAN 596
D 637 TGRINDTKYQAEARNRGELVSSVOCYMRNPECTEEBALSHTVYGIINALKELNWEIAN 596
Db 494 FIRLVNDLTSHTVQORGHVASCIDCYNNQH-GVSKBAVKILQWATDCWKEINEE--- 349
QY 697 PASNAPLCVR-----LLENTARVQWLFMYRVDGFGISDKEMKHVSRTLPDPA 746
D 697 PASNAPLCVR-----LLENTARVQWLFMYRVDGFGISDKEMKHVSRTLPDPA 746
Db 550 -----CMROSQSVGHLMRIVNLARLTDVSYKYGDGY-TDSQQLKQFVKGLFVDPI 630
QY 637 TGRINDTKYQAEARNRGELVSSVOCYMRNPECTEEBALSHTVYGIINALKELNWEIAN 596
D 637 TGRINDTKYQAEARNRGELVSSVOCYMRNPECTEEBALSHTVYGIINALKELNWEIAN 596
Db 494 FIRLVNDLTSHTVQORGHVASCIDCYNNQH-GVSKBAVKILQWATDCWKEINEE--- 349
QY 697 PASNAPLCVR-----LLENTARVQWLFMYRVDGFGISDKEMKHVSRTLPDPA 746
D 697 PASNAPLCVR-----LLENTARVQWLFMYRVDGFGISDKEMKHVSRTLPDPA 746
Db 550 -----CMROSQSVGHLMRIVNLARLTDVSYKYGDGY-TDSQQLKQFVKGLFVDPI 630
RESULT 14
Y4KT RHISN
ID Y4KT RHISN STANDARD; PRT; 516 AA.
AC P5538;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 64.4 kDa protein Y4KT.
GN Y4KT.
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: STRONG, TO B.JAPONICUM HYPOTHETICAL 55.4 kDa PROTEIN
CC IN CYP117 3' REGION (ORF7).
CC -1- SIMILARITY: Contains 2 PFTB repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000082; AAB91751.1; -.
DR PIR; T10874; T10874.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR008930; Terp cyc toroid.
DR Pfam; PF00432; prenyltrans; 2.
KW Hypothetical protein; Repeat; Plasmid.
FT REPEAT 45 86 PFTB 1.
FT REPEAT 401 443 PFTB 2.
SQ SEQUENCE 516 AA; 55108 MW; E4BD8C45PB52DCEB CRC64;
Query Match 5.5%; Score 219.5; DB 1; Length 516;
Best Local Similarity 24.2%; Pred.No. 7.8e-08;
Matches 113; Conservative 48; Mismatches 176; Indels 129; Gaps 21;
QY 1 MFQSMGD-GETNPSTAYDTAWARIPSIDSGAPQPPQTQLMNNQLPDGSGWGECEIFLA 39
D 1 MFQSMGD-GETNPSTAYDTAWARIPSIDSGAPQPPQTQLMNNQLPDGSGWGECEIFLA 39
Db 15 LLEMSDGSVGSVYDTRALQF-----GGNVTRQDAYAWLLAQOQADGGWSADFLF 70
QY 60 YDRVLNTLACLLTKIWNKGD-----IQVKGVEFVRKHEEMKDEADNHRPSGFVVF 113
D 60 YDRVLNTLACLLTKIWNKGD-----IQVKGVEFVRKHEEMKDEADNHRPSGFVVF 113
Db 71 --RHAPTAALLAQ---RADPLPGAADVAQATFLERQADPYAHAVPEDAPIGAELIL 125
QY 114 PAMLDKSL--GLDLPVLPFISQIHKQKQKLOKI-----PLNVLHNHQTALLYSL 164
D 114 PAMLDKSL--GLDLPVLPFISQIHKQKQKLOKI-----PLNVLHNHQTALLYSL 164
Db 126 PQLCGEASLLGGVAFPRH--PALPLRQACLVLGAVATLPFGHPL--LHGSWEAWGTPT 182
QY 165 EGLQDVVDWQBITNLQSDGSGFLSPASTACVFVMT-----QMKRCLHFL-----NFVL 213
D 165 EGLQDVVDWQBITNLQSDGSGFLSPASTACVFVMT-----QMKRCLHFL-----NFVL 213
Db 183 AACPD-----DDSGIGISPAATAWAHAHVTQSTPQVGRADAYLQASRATR 230
```

```
QY 214 SKFGDYVPCHYPLDLFERLWAVDTVERLGIDRYFKKEIKESLDYVRYWDAB---BGVCH 270
D 214 SKFGDYVPCHYPLDLFERLWAVDTVERLGIDRYFKKEIKESLDYVRYWDAB---BGVCH 270
Db 231 SGIEGVWVNVWPIVFPFCWCSYTLHLAAGL--FAHPALDEAVRVIVVAQLDARLGVRLGP 288
QY 271 ARCNPIPDVDDTAMGLRLRLHGVNVSSDVLENFRDEKGDFFCFAGQTQIGVTDLNLNLYR 330
D 271 ARCNPIPDVDDTAMGLRLRLHGVNVSSDVLENFRDEKGDFFCFAGQTQIGVTDLNLNLYR 330
Db 289 A-LHPAADADDTAVALCVLRLAGRPVADALRHP--EIGELP----- 327
QY 331 CSQVCFPEKIMEEAKTFTTNLQNAL-----AKNNAFDKWAYKXDLPG 375
D 331 CSQVCFPEKIMEEAKTFTTNLQNAL-----AKNNAFDKWAYKXDLPG 375
Db 328 --VTFPGER---NASVSTNIHALRLGLKPAAGTSAYVEANRNPGLWLDNEK----- 376
QY 376 VEYAIKYPWHRSMPLREARSYIEQFGSNDVWLKTYVKMLYSNEK 421
D 376 VEYAIKYPWHRSMPLREARSYIEQFGSNDVWLKTYVKMLYSNEK 421
Db 377 -----WHVS-----WLYPTAHAAVAALAQGK 396
RESULT 15
YL49 BRAJA
ID YL49 BRAJA STANDARD; PRT; 516 AA.
AC Q45221;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein blr2149 (ORF7).
GN BLR2149.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
RA Tully R.E., Keister D.L.;
RT "Cloning and mutagenesis of a cytochrome P-450 locus from
RT Bradyrhizobium japonicum that is expressed anaerobically and
RT symbiotically.";
RL Appl. Environ. Microbiol. 59:4136-4142(1993).
RN [2]
SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
RA Tully R.E., van Berkum P., Lovins K.W., Keister D.L.;
RT "Identification and sequencing of a cytochrome P450 gene cluster from
RT Bradyrhizobium japonicum.";
RL Biochim. Biophys. Acta 1398:243-255(1998).
RN [3]
SEQUENCE FROM N.A.
RX STRAIN=USDA 110;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL Nat. Res. 9:189-197(2002).
CC -1- SIMILARITY: Contains 2 PFTB repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U12678; AAC28895.1; -.
DR EMBL; AP005942; BAC47414.1; ALT_INIT.
DR PIR; I40214; I40214.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR008930; Terp cyc toroid.
DR Pfam; PF00432; prenyltrans; 2.
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 17, 2004, 20:56:44 ; Search time 40.0877 Seconds  
(without alignment)  
5871.540 Million cell updates/sec

Title: US-10-041-007-39  
Perfect score: 3972  
Sequence: 1 MFQSMGDTNPSAYDTAWV.....GISDKEMKHVSRFLFDPA 746

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3972	100.0	873	10 Q947C4	Q947C4 ginkgo bilo
2	2735.5	68.9	868	10 Q38710	Q38710 abies grand
3	2733.5	68.3	853	10 Q94FW1	Q94FW1 abies grand
4	2050.5	51.6	782	10 Q9SAU6	Q9SAU6 abies grand
5	2050.5	51.6	817	10 Q81086	Q81086 abies grand
6	2040.5	51.4	816	10 Q94FW2	Q94FW2 abies grand
7	1382	34.8	787	10 Q22667	Q22667 stevia reba
8	1373	34.6	800	10 Q9ST35	Q9ST35 lycopersico
9	1333	33.6	829	10 Q9MAX2	Q9MAX2 croton subl
10	1327	33.4	827	10 Q9LRE1	Q9LRE1 scoparia du
11	1320	33.2	802	10 Q38802	Q38802 arabidopsis
12	1309	33.0	799	10 Q9FV9	Q9FV9 lactuca sat
13	1296.5	32.6	823	10 Q9ZTN8	Q9ZTN8 cucurbita m
14	1294.5	32.6	822	10 Q84LM0	Q84LM0 cucurbita m
15	1293.5	32.6	823	10 Q41771	Q41771 zea mays (m
16	1286.5	32.4	827	10 Q9ZTN7	Q9ZTN7 cucurbita m

17	1275.5	32.1	801	10	O04408	O04408 pisum sativ
18	1257.5	31.7	577	10	Q94KA3	Q94KA3 picea abies
19	1252	31.5	581	10	O64404	O64404 abies grand
20	1244	31.3	579	10	Q94FW3	Q94FW3 abies grand
21	1215.5	30.6	630	10	Q9M7D1	Q9M7D1 abies grand
22	1209.5	30.5	634	10	Q94KA5	Q94KA5 picea abies
23	1203	30.3	637	10	Q94FV9	Q94FV9 abies grand
24	1193	30.0	637	10	Q9M7C9	Q9M7C9 abies grand
25	1182	29.8	629	10	Q84KL6	Q84KL6 pinus taeda
26	1181.5	29.7	633	10	Q94KA4	Q94KA4 picea abies
27	1169	29.4	628	10	O84KL3	O84KL3 pinus taeda
28	1151.5	29.0	593	10	O64405	O64405 abies grand
29	1139.5	28.7	623	10	Q94FW0	Q94FW0 abies grand
30	1134.5	28.6	785	10	Q9FRX5	Q9FRX5 cumulus sat
31	1131	28.5	615	10	Q84KL2	Q84KL2 pinus taeda
32	1131	28.5	618	10	Q948Z0	Q948Z0 abies grand
33	1130	28.4	685	10	Q7XNG0	Q7XNG0 oryza sativ
34	1126.5	28.4	627	10	Q84SM8	Q84SM8 picea abies
35	1122	28.2	788	10	Q9FXV8	Q9FXV8 lactuca sat
36	1116	28.1	630	10	Q9M7D0	Q9M7D0 abies grand
37	1106.5	27.9	627	10	Q84KL4	Q84KL4 pinus taeda
38	1094.5	27.6	574	10	Q84KL5	Q84KL5 pinus taeda
39	1072	27.0	784	10	Q9XEH9	Q9XEH9 stevia reba
40	1066.5	26.9	789	10	Q39548	Q39548 cucurbita m
41	1054	26.5	784	10	Q9XEI0	Q9XEI0 stevia reba
42	1025.5	25.8	785	10	O64952	O64952 arabidopsis
43	1025.5	25.8	785	10	Q9SAK2	Q9SAK2 arabidopsis
44	961.5	24.2	830	10	Q852S2	Q852S2 oryza sativ
45	843	21.2	683	10	Q7XLE0	Q7XLE0 oryza sativ

#### ALIGNMENTS

##### RESULT 1

Q947C4  
ID Q947C4 PRELIMINARY; PRT; 873 AA.  
AC Q947C4;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Levopimaradiene synthase.  
OS Ginkgo biloba (Ginkgo).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.  
OX NCBI\_TaxID=3311;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21381393; PubMed=11488601;  
RA Schepmann H.G.; Pang J.; Matsuda S.P.T.;  
RT "Cloning and characterization of Ginkgo biloba levopimaradiene  
RT biosynthesis which catalyzes the first committed step in ginkgolide  
RT biosynthesis."  
RL Arch. Biochem. Biophys. 392:263-269(2001).  
DR EMBL; AF331704; AAC0965.1; -;  
DR GO; GO:0016829; P:metabolism; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR005630; Terpene synth C.  
DR InterPro; IPR008949; Terpenoid synth.  
DR InterPro; IPR008930; Terp\_cyc\_Teroid.  
DR InterPro; IPR001906; Terp\_synth-like.  
DR Pfam; PF01397; Terpene synth; 1.  
DR Pfam; PF03936; Terpene synth C; 1.  
SQ SEQUENCE 873 AA; 100290 MW; F2C980BB19F9931E CRC64;

Query Match 100.0%; Score 3972; DB 10; Length 873;  
Best Local Similarity 100.0%; Pred. No. 8.3e-283;  
Matches 746; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFQSMGDTNPSAYDTAWVIPSIDGSGAPFPOTLQWLNQLPDGSGWGEICIFLAY 60  
|||||  
DB 128 MFQSMGDTNPSAYDTAWVIPSIDGSGAPFPOTLQWLNQLPDGSGWGEICIFLAY 187

QY 61 DRVLNTLACLLTLKIWNKGDIOVQGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEA 120  
 DB 188 DRVLNTLACLLTLKIWNKGDIOVQGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEA 247  
 QY 121 KSLGLDLPYHLPPISQIHQKQKLOKIPLNVLHNTALLYSLEGLQDVVDWQEIITNLQ 180  
 DB 248 KSLGLDLPYHLPPISQIHQKQKLOKIPLNVLHNTALLYSLEGLQDVVDWQEIITNLQ 307  
 QY 181 SRDGSFLSPASTACVFMHTQNKRCCLHFLNVLKSGDYVPCHPYLDLFLERLWAVDIVER 240  
 DB 308 SRDGSFLSPASTACVFMHTQNKRCCLHFLNVLKSGDYVPCHPYLDLFLERLWAVDIVER 367  
 QY 241 LGIDRYFKEIKESLDYVYRYWDAERGVCWARGNCPIDVDVDTAMGLRIILRLHGYNVSSDV 300  
 DB 368 LGIDRYFKEIKESLDYVYRYWDAERGVCWARGNCPIDVDVDTAMGLRIILRLHGYNVSSDV 427  
 QY 301 LENFRDEKDFCFAGQTOIGVTDNLNLYRCQVCPGGEKIMEEAKTFTTNHLQNALAKN 360  
 DB 428 LENFRDEKDFCFAGQTOIGVTDNLNLYRCQVCPGGEKIMEEAKTFTTNHLQNALAKN 487  
 QY 361 NAFDKWAVKDLPGVEYEAIKYPWHRSMRPLEARSYIEQFGSNDVWLKTVYKMLVYSNE 420  
 DB 488 NAFDKWAVKDLPGVEYEAIKYPWHRSMRPLEARSYIEQFGSNDVWLKTVYKMLVYSNE 547  
 QY 421 KYLELAKLDFNMVQALHOKETQIHVSWWRSGFNDLFTTRQRPVEMYFSAVSMFPEPA 480  
 DB 548 KYLELAKLDFNMVQALHOKETQIHVSWWRSGFNDLFTTRQRPVEMYFSAVSMFPEPA 607  
 QY 481 ACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLY 540  
 DB 608 ACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLY 667  
 QY 541 NTVNGFGKDLKEQGRDVLGYLRKWEGLLASYTKEAENSAKYVPTFNEYENAKVSTA 600  
 DB 668 NTVNGFGKDLKEQGRDVLGYLRKWEGLLASYTKEAENSAKYVPTFNEYENAKVSTA 727  
 QY 601 LATVNLNSIFFTGELLPDYILOQVDLRSKFLHLVSLTGRINDTKYQAEARNRGLVSSV 660  
 DB 728 LATVNLNSIFFTGELLPDYILOQVDLRSKFLHLVSLTGRINDTKYQAEARNRGLVSSV 787  
 QY 661 QCYMRENPECTEERLASHVYGIIDNALKELNWLANPASNAPLCVRRLLFNTARVMOLF 720  
 DB 788 QCYMRENPECTEERLASHVYGIIDNALKELNWLANPASNAPLCVRRLLFNTARVMOLF 847  
 QY 721 MYRDGFGISDKEMKOHVSRTPDPVA 746  
 DB 848 MYRDGFGISDKEMKOHVSRTPDPVA 873

## RESULT 2

Q38710 ID Q38710 PRELIMINARY; PRT; 868 AA.  
 AC Q38710;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Abietadiene cyclase.  
 GN AC22.  
 OS Abies grandis (Grand fir).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.  
 OX NCBI\_TaxID=46611;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Stem;  
 RA Stoffer-Vogel B., Wildung M.R., Vogel G., Croteau R.B.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U50768; AAB05407.1; --  
 DR HSSP; Q40577; SEAS.  
 DR GO; GO:0016629; F1yase activity; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001330; Prenyltrans.  
 DR InterPro; IPR005630; Terpene\_synth\_C.

DR InterPro; IPR008949; Terpenoid synth.  
 DR InterPro; IPR008930; Terp\_cyc\_toroid.  
 DR InterPro; IPR001906; Terp\_synth-like.  
 DR Pfam; PF00432; prenyltrans; 1.  
 DR Pfam; PF01397; Terpene\_synth; 1.  
 DR Pfam; PF03936; Terpene\_synth\_C; 1.  
 SQ SEQUENCE 868 AA; 99536 MW; AD5E79F56B70D25C CRC64;

Query Match 68.9%; Score 2735.5; DB 10; Length 868;  
 Best Local Similarity 67.8%; Pred. No. 5.2e-192;

Matches 505; Conservative 103; Mismatches 134; Indels 5; Gaps 4;

QY 1 MFQSMGDMGTNSAYDTAWVARIPSDGSGAPOPFOTLOWLNNQPLPDGSGWGECEIFLAY 60  
 DB 126 MFRCMGYGETNPSAYDTAWVARIPAVDGSNDPHFETVEWILQNLKDGSGWGECEIFLAY 185  
 QY 61 DRVLNTLACLLTLKIWNKGDIOVQGVFVRKHMEEMKDEADNHRPSGFEVFPAMLDEA 120  
 DB 186 DRILATLACIIITLWRTGETQVQKGIETPRTQAGKMEDEADSHRPSGFEVFPAMLKEA 245  
 QY 121 KSLGLDLPYHLPPISQIHQKQKLOKIPLNVLHNTALLYSLEGLQDVVDWQEIITNLQ 180  
 DB 246 KILGLDLPYDLPPKLIIEKREAKLARIPTDVLVYALPTLLYLSLEGLQEIIVDKIMKQJ 305  
 QY 181 SRDGSFLSPASTACVFMHTQNKRCCLHFLNVLKSGDYVPCHPYLDLFLERLWAVDIVER 240  
 DB 306 SKDGSFLSPASTAAVFMRTGNKKCLDFNLVLFKFGNHVPCHYPLDLFLERLWAVDIVER 365  
 QY 241 LGIDRYFKEIKESLDYVYRYWDAERGVCWARGNCPIDVDVDTAMGLRIILRLHGYNVSSDV 300  
 DB 366 LGIDRYFKEIKESLDYVYSHWD- ERGIGWARENPVPDI DDTAMGLRIILRLHGYNVSSDV 424  
 QY 301 LENFRDEKDFCFAGQTOIGVTDNLNLYRCQVCPGGEKIMEEAKTFTTNHLQNALAKN 360  
 DB 425 LKTRFDENGEFCFLGQTOIGVTDNLNLYRCQVCPGGEKIMEEAKTFTTNHLQNALAKN 484  
 QY 361 NAFDKWAVKDLPGVEYEAIKYPWHRSMRPLEARSYIEQFGSNDVWLKTVYKMLVYSNE 420  
 DB 485 DAFDKWAFKKNIRGEVEYALKYPMHKSMPRLEARSYIENYGPDDVWLKTVYKMLVYSNE 544  
 QY 421 KYLELAKLDFNMVQALHOKETQIHVSWWRSGFNDLFTTRQRPVEMYFSAVSMFPEPA 480  
 DB 545 KYLELAKLDFNMVQALHOKETQIHVSWWRSGFNDLFTTRQRPVEMYFSAVSMFPEPA 604  
 QY 481 ACRIAYAKTSCLAVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLY 540  
 DB 605 KCREVYTKTSNFTVILDDLYDTHGSLDLDLKFSEAVRRWDISVLDSVRDNLKVCFLGLY 663  
 QY 541 NTVNGFGKDLKEQGRDVLGYLRKWEGLLASYTKEAENSAKYVPTFNEYENAKVSTA 600  
 DB 664 NTFNDIAKEGREGQGRDVLGYIQNVKVKQLAEYATKEAENSAKYVPSFNEYENAKVSTA 723  
 QY 601 LATVNLNSIFFTGELLPDYILOQVDLRSKFLHLVSLTGRINDTKYQAEARNRGLVSSV 660  
 DB 724 LGTVVLISALFTGEVLTDEVLSDKIDRESRFLQMLGLTGRVNDTKYQAEARGOGEVASAI 783  
 QY 661 QCYMRENPECTEERLASHVYGIIDNALKELNWLANPASNAPLCVRRLLFNTARVMOLF 720  
 DB 784 QCYMRENPECTEERLASHVYGIIDNALKELNWLANPASNAPLCVRRLLFNTARVMOLF 841  
 QY 721 MYRDGFGISDKEMKOHVSRTPDPVA 746  
 DB 842 MQGDGLTLSDHMEIKHVKNCFLFPVA 868

## RESULT 3

Q94FW1 ID Q94FW1 PRELIMINARY; PRT; 853 AA.  
 AC Q94FW1;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Abietadiene synthase (fragment).

[illegible]

Qy	423	LELAKLDFNMVQALHQKETOHI	VSWRRSGGNDLFTFTQRPVEM	YFVSVAVSWFEPFAAC	488
Db	457	LELAKLDFNIIQSIHQEEMKNVT	SWFRDSGLPLFTFABERPLIE	FYFLVAAGTYEPQYAKC	516
Qy	483	RIAYAKTSCLAIVLDLVDYTHGS	LDLKLKFSFAVRWDI	SVLDSVRDNLQKVCFLGLYNT	542
Db	517	RFLFTKVCACLQTVLDDMYDTT	GTGDELKLFTEAVRWDLSFTEN	LPD-YMKLCYQIYYDI	575
Qy	543	VNGFGDKLKEQGRDVLGYLRKW	WEGLLASVTKAEWSAAKYVPT	FNEVVENAKVSIALA	602
Db	576	VHEVAWEAKEQGRSELVSFRK	GWEDYLLGYEEAEWLAEEVPL	DEYKNGITSIGOR	635
Qy	603	TVVLNSI-FTTGELLDPYILQOV	DL--RSKFLHLVSLTGR	LINDTKTYOAEKRNRELVS	659
Db	636	ILLLSGVLMGQLLSQBALEKVD	YPGRRVLTELSLSRLADDTK	TYKAERKARGELASS	695
Qy	660	VOCYMRNPCTEEREAALSHVGI	IDNALKENWELANPASNA	PLCVRRLLENTAVMWOLF	719
Db	696	IECYMKHPECTEEREAALHII	YSILEPAVKELTRFLKP-D	DDVPFACKMPLPEETRTWVI	754
Qy	720	YMYRDGFGISDKEMKHVSRIT	FLDFV 745		
Db	755	PKDGDGFGVSKLEVKDHIKEC	LIEPL 780		
RESULT 5					
ID	081086	PRELIMINARY; PRT; 817 AA.			
ID	AC	081086;			
DT	01-NOV-1998	(TRENBLRel. 08, Created)			
DT	01-NOV-1998	(TRENBLRel. 08, Last annotation update)			
DT	01-OCT-2003	(TRENBLRel. 25, Last annotation update)			
DE	E-alpha-bisabolene synthase.				
DS	Abies grandis (Grand fir).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.				
OX	NCBI_TaxId=46611;				
FN	[1]				
RP	SEQUENCE FROM N.A.				
RA	MEDLINE=98284004; PubMed=9618485;				
RA	Bohlmann J., Crock J., Jetter R., Croteau R.;				
RT	"Terpenoid-based defenses in conifers: cDNA cloning, characterization,				
RT	and functional expression of wound-inducible (E)-alpha-bisabolene				
RT	synthase from grand fir (Abies grandis).";				
RL	Proc. Natl. Acad. Sci. U.S.A. 95:6756-6761(1998).				
DR	EMBL; AF006195; AAC34192.1; -.				
DR	HSSP; Q40577; SEAS.				
DR	GO; GO:0016829; P:lyase activity; IEA.				
DR	GO; GO:0008152; P:metabolism; IEA.				
DR	InterPro; IPR005830; Terpene synth_C.				
DR	InterPro; IPR008949; Terpenoid synth.				
DR	InterPro; IPR008930; Terp_cyc toroid.				
DR	InterPro; IPR001906; Terp synth-like.				
DR	Pfam; PF01397; Terpene synth; 2.				
DR	Pfam; PF03936; Terpene synth_C; 1.				
SQ	SEQUENCE 817 AA; 93749 MW; 95FB06D0BC0DE1B4B CRC64;				
Query Match					
Best Local Similarity 51.6%; Score 2050.5; DB 10; Length 817;					
Matches 379; Conservative 152; Mismatch 206; Indels 9; Gaps 7					
Qy	6	GDGET--NPSAYDTAWARI	SIDCSGAPQPPQTQLWINN	LPDGSWGECIELAYDRV	63
Db	73	GDGESMITPSAYDTAWARV	PAIDSGARPPQPTVDWILK	NLDGSGWIGSHFLLSDRL	132
Qy	64	LNTLACLITLTKIWNKGD	IQVQKGVFVRKHEEMKDEAD-N	HRPSPGFVVPFAMLDEAKS	122
Db	133	LATLSCLVLLKWNVGD	LQVEGIEFIKSNLELVKDETD	QDSLSVTDFFIIFSLREAGS	192
Qy	123	LGLDLPYHLPTFSIQHKRQK	LKQIKPIPLVNLHNQTALLYS	LEGQDVVDWQETNQLQR	182
Db	193	LRLGLPYDLPYHLTQTRER	LAKLSREEIYAVPSPLYSLEG	QDI VEWERIMEVQSQ	252

Qy	183	DGSEFLSSPASTACVPMFTQNKRCUHLFLNPNVLVSKGDFYPCHPYPLDI.FERLMAVDVTVERLG	243
Db	253	DGSEFLSSPASTACVPMFTGDAKCLLEFLNSVMIKFGNFVPCLYPDVLLERLLVDNIVRLG	312
Qy	243	IDRVFKKEIKESLDVYRYWDAERGVGWARCNPIPDVDDTAMGLRILRLHGYNVSSDVLB	302
Db	313	IYRHFEKEIKALDYVYRHWN-ERGIHGRNLNPDIADLETTALGFRLRLHRYNVSPAIPD	371
Qy	303	NFRDEKGDFFCFAGOTQIGVTDNINLNYRCSQVCFGEKIMEBAKTTTTHLQNALAKNNA	362
Db	372	NFKDANGKFCSTGQFNKDVASMLNLYRASQLAPGENILDEAKSFATKYLREALEKSET	431
Qy	363	FDKAVVAKDLPGEEVYAIKYPHRSMPRLARSYTEQFGSNDVWLKGTYYKMLYVSNEKY	422
Db	432	SSAWNKNQNSQETKYALKTSWHASVPRVEAKRYCQVYRPDVARIAKCVIKLPYVNNKXP	491
Qy	423	LELAKLDFNMVQALHQKETQHVSVNRWBSGENDLFTTRQRPVEMYFSVAVSMFEPEFAAC	482
Db	492	LELKLDFNIIQSIOHEEMKNVTSWFRDGLPLTFARERPLEFYFLVAAGTVEPOYAKC	551
Qy	483	RIAVAKTSLAVIIDLDYTHGSLDLDLKFSEAVRRWDISVLDSVRDNOLKVCFLGLYNT	542
Db	552	RFLFTKVCACLTGLVDDMYDTYGTLDLKLFTFAVRRWDLSTFENTLPD-YMKLCYQIYYDI	610
Qy	543	VNGFGDKLQKQGRDVLGYSKVMWEGLLASTYKAEASAAKXVPTFNEYVENAKVSIALA	602
Db	611	VHEVAWEAKEQGRELVSFFPKGWEDYLLGYEEAEWLAAEVPTLDEVIKNGITSIGQR	670
Qy	603	TVLVNSI-FTTGELLDPYILOQVDL--RSKFLHLVSLTGRLLINDTKYQAEARNRGLVSS	659
Db	671	ILLSSGLVMDQLLSQELAEKVDYDGGRRVLTETELNSLISRLADDTKTYKAEARGELASS	730
Qy	660	VOCYMRNPECTEREALSHVVGIIIDNALKEINWELANPASNAPLCVRRLLENTARVMQLF	719
Db	731	IECYMKHPECTEREALDHIYSILEPAVKELTREFLKP-DDVFPACKGMFLFEETRVTWVI	789
Qy	720	YMYRDGFGISDKEMKHDSVRTLPDVP 745	
Db	790	FKDGGDGVSKLVKDHKECLIEPL 815	
RESULT 6			
AC	Q94FW2	PRELIMINARY;	PRT; 816 AA.
ID	Q94FW2		
AD	Q94FW2		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	E-alpha-bisabolene synthase (Fragment).		
OS	Abies grandis (Grand fir).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Abies.		
OX	NCBI_taxid=46611;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21297238; PubMed=11404343;		
RA	Trapp S.C., Croteau R.B.;		
RT	"Genomic organization of plant terpene synthases and molecular		
RT	evolutionary implications";		
FL	Genetics 158, 811-832(2001).		
DR	EMBL; AF346515; AAK83562.1;		
DR	GO; GO:0016829; F-lyase activity; IEA.		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR005630; Terpene synth C.		
DR	InterPro; IPR008949; Terpenoid synth.		
DR	InterPro; IPR008930; Terp cyc toroid.		
DR	InterPro; IPR001906; Terp_synth-like.		
DR	Pfam; PF01397; Terpene synth_2.		
DR	Pfam; PF03936; Terpene synth_C. 1.		
FT	NON TER		
SQ	SEQUENCE 816 AA		
FT	SEQUENCE 816 AA; 684FDS28D1EC67E4 CRC64;		

```

Query Match      51.4%; Score 2040.5; DB 10; Length 816;
Best Local Similarity 50.5%; Pred. No. 5.1e-141;
Matches 377; Conservative 153; Mismatches 207; Indels 9; Gaps 7;

QY 6 GDGET--NPSAYDTAWARIPSIDSGAPQPPQTQIQLWILNNQPDGSGWEECIFIPLAYDRV 63
DB 72 GDGSMITPSAYDTAWARIPSIDSGAPQPPQTQIQLWILNNQPDGSGWEECIFIPLAYDRV 131
QY 64 LNTIACLTTLKWNKGDIOVQKGVFVRKHEEMKDEAD-NHRPSGFEVVPVPPAMLDEAKS 122
DB 132 LATLSCLVLLKWNVGDIOVQKGVFVRKHEEMKDEAD-NHRPSGFEVVPVPPAMLDEAKS 191
QY 123 LGLDLPVHLPTSIHQKROKQKLOKIPLVNLHHTALLYSLEGIQDVVDVQOEITNLQSR 182
DB 192 LRLGLPYDLPVHLPTSIHQKROKQKLOKIPLVNLHHTALLYSLEGIQDVVDVQOEITNLQSR 251
QY 183 DGSFLSSPASTACVFMHTQNRKHLFLNVLKSGFYVPCHPYPLDLFERLMAVDVVRBLG 242
DB 252 DGSFLSSPASTACVFMHTQNRKHLFLNVLKSGFYVPCHPYPLDLFERLMAVDVVRBLG 311
QY 243 IDRVFKKEIKESLDVYRYWDAERGCVGWARCNPIPDVDDTAMGLRILRLHGVNYSDDVLE 302
DB 312 IYRHFKEIKESLDVYRYWDAERGCVGWARCNPIPDVDDTAMGLRILRLHGVNYSDDVLE 370
QY 303 NFRDEKDPFCFAGQTOIGVTDNLNLYRSCVQCPGKIMEEAKTFTTNHLQNALAKNA 362
DB 371 NFDKANGKFCSTQGNKNDVASMNLRYASQAPFENILDEAKSFATKYLREALEKSET 430
QY 363 FDKWAKKDLPGVEYAIKYPMHRSMPRLARSYIEQFGSNDVWLGKTYVMKLYVSNKY 422
DB 431 SSANNKQNLQSEIKYALKTSWASVPRVAKRYCQVTRPDYARAKCYKLPYVNNKXF 490
QY 423 LELAKLDPNMVQALHOKETQHIWWSRGSGFNDLFTTRQRPVEMYFSAVMSFEPFAAC 482
DB 491 LELAKLDPNMVQALHOKETQHIWWSRGSGFNDLFTTRQRPVEMYFSAVMSFEPFAAC 550
QY 483 RIAYAKTSLAVILDDLYTHGSLDDKLKFEAVRRWDISVLDSVRDNLQKVCPLGLYNT 542
DB 551 RPLFTTKVACLQVLDMDYDTGTDELKLFTEAVRRWDVSVFENLPD-YMKLCYQIYYDI 509
QY 543 VNGFGKDLQKQGRDVLGKRWEGILLASYTKAEWSAAKVPFTFNEYVENAKYSIALA 502
DB 610 VHEVAWEAKQGRHLSVFPKRGWEDYLLGYEEAEWLAAYVPSLDHYKNGITSIGOR 569
QY 603 TVVLNSI--PFTGELLPDYILQOVDL--RSKFLHLVSLTGRINDTKTYOAEARNRGELVSS 559
DB 670 ILLSGVLIMDQQLSQALEKVDYVGRRLVTELSRLSLRADDTKTYKAERKARGELASS 729
QY 660 VQCYMRENPCTEEREAALSHVGIIDNALKELNWLANPASNAPLCVRRLLFTWARVMOLF 719
DB 730 IECYMKDHPCTEEREAALSHVGIIDNALKELNWLANPASNAPLCVRRLLFTWARVMOLF 788
QY 720 YMYRDGSGISDKMKDHSRTLPDPV 745
DB 789 FROGDGFGVSKLEVKDHIKECLIEPT 814

RESULT 7
O22667 PRELIMINARY; PRT; 787 AA.
AC O22667;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Copalyl pyrophosphate synthase.
GN CP881.
OS Stevia rebaudiana (Stevia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Asteroidae; Eupatoriaceae; Stevia.
OX NCBI_TaxID=55670;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Leaf;
RX MEDLINE=99435880; PubMed=10504563;
RA Richman A.S., Gilzen M., Starratt A.N., Yang Z., Brandle J.E.;
RT "Diterpene synthesis in Stevia rebaudiana: recruitment and up-
regulation of key enzymes from the gibberellin biosynthetic pathway.";
RL Plant J. 19:411-421(1999).
DR EMBL; AF034545; BAB87091.1; -.
DR GO; GO:0016825; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpenoid synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene_synth; 1.
DR Pfam; PF03936; Terpene_synth_C; 1.
SQ SEQUENCE 787 AA; 90445 MW; 224575732606DCA1 CRC64;

Query Match      34.8%; Score 1382; DB 10; Length 787;
Best Local Similarity 42.7%; Pred. No. 1.1e-92;
Matches 280; Conservative 122; Mismatches 235; Indels 18; Gaps 9;

QY 1 MFQSMGDSGTNPSTAYDTAWARIPSIDSGAPQPPQTQIQLWILNNQPDGSGWEECIFIPLAY 60
DB 98 MFGSMNDGSEINVSAYDTAWARIPSIDSGAPQPPQTQIQLWILNNQPDGSGWEECIFIPLAY 157
QY 61 DRVLNTLACLTTLKWNKGDIOVQKGVFVRKHEEMKDEAD-NHRPSGFEVVPVPPAMLDEA 120
DB 158 DRIINTLACVIALTSWNVHPSKCEKGLNFKLENICKLEDAENAEHMPGFEVVPVPPAMLDEA 217
QY 121 KSLGLDLPVHLPTSIHQKROKQKLOKIPLVNLHHTALLYSLEGIQDVVDVQOEITNLQ 180
DB 218 KGLNIEVEDPTALKEIYARRDKLTKIPVNLHHTALLYSLEGIQDVVDVQOEITNLQ 276
QY 181 SRDGSFLSSPASTACVFMHTQNRKHLFLNVLKSGFYVPCHPYPLDLFERLMAVDVVRBLG 240
DB 277 CKDGSFLSSPASTACVFMHTQNRKHLFLNVLKSGFYVPCHPYPLDLFERLMAVDVVRBLG 336
QY 241 LGIDRYFKKEIKESLDVYRYWDAERGCVGWARCNPIPDVDDTAMGLRILRLHGVNYSDDV 300
DB 337 LGIARYFKKEIKESLDVYRYWDAERGCVGWARCNPIPDVDDTAMGLRILRLHGVNYSDDV 395
QY 301 LENFRDEKDPFCFAGQTOIGVTDNLNLYRSCVQCPGKIMEEAKTFTTNHLQNALAKN 360
DB 396 FRQF-EKQKGFVCFAGQTOIGVTDNLNLYRSCVQCPGKIMEEAKTFTTNHLQNALAKN 454
QY 361 NAFDQWAKKDLPGVEYAIKYPMHRSMPRLARSYIEQFGSNDVWLGKTYVMKLYVSN 419
DB 455 ELLDKWIIAKOLPGVEYAIKYPMHRSMPRLARSYIEQFGSNDVWLGKTYVMKLYVSN 514
QY 420 EYKLELAKLDPNMVQALHOKETQHIWWSRGSGFNDLFTTRQRPVEMYFSAVMSFEPFAAC 478
DB 515 NYLEMAKLDYNNYVAVLQEWYTIQQWYVDIGIEKFPESDNKSVLSVSYLAASIFEPE 574
QY 479 PAACRIAYAKTSLAVILDDLYTHGSLDDKLKFEAVRRWDISVLDSVRDNLQKVCPL 537
DB 575 RSKERIAWAKTTLVNDKITSIFDSQSSEKEDITAFIDKFRNKSSEKKSXGNGEPMHEVMV 634
QY 538 GLYNTVNGFGKDLQKQGRDVLGKRWEGILLASYTKAEWSAAKVPFTFNEYVENAKV 597
DB 635 ALKKTILHGFALDAMTHSQDIHPQLHQAEMWMLTKLQDGVVDVTAELMVQMIN-WTAGRWV 693
QY 598 SIATAT-----VVLNSIIFTGELLPDYILQOVDL--RSKFLHLVSLTGRINDTKTYOAEARNRGELVSS 644
DB 694 SKELLTHPOYQRLSTVNTSVCHDITKLNHFKENSTTVDSKVQLVQL---VPSDT 745

RESULT 8
Q9ST35 PRELIMINARY; PRT; 800 AA.
AC Q9ST35;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

```

```
DE Copalyl diphosphate synthase.
GN CPS.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RP SEQUENCE FROM N.A.
RA Imai R.;
RT "Tomato cDNA for copalyl diphosphate synthase.";
RL EMBL; A8015675; BAA84918.1; -.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth. 1.
DR Pfam; PF03936; Terpene synth. C; 1.
SQ SEQUENCE 800 AA; 91698 MW; 040B522FE19902F9 CRC64;

Query Match 34.6%; Score 1373; DB 10; Length 800;
Best Local Similarity 39.5%; Pred. No. 5.2e-92;
Matches 298; Conservative 132; Mismatches 257; Indels 68; Gaps 18;

QY 1 MFQSGMDGNETPSAYTAWVARIPIIDGSGAPQPTLOWLNNQLPDGSGWBECEIFLAY 60
DB 103 MLGSMGDGEISVSAYTAWVAMVDKVTETPQPSLEWIANQALDAGWDGNSIFLAV 162
QY 61 DRVINTLACLLTLKINWKGDIQVQGVFVRKHEEMKDEADNHRSGFVVPAMLEA 120
DB 163 DRVINTLACVTLKSNWLNHPDKILLGSMFRENLSRIGDENAEHMPIGFEVAPSLIEA 222
QY 121 KSLGLDLPVHLPIISOIHQKROKLOKIPLVNLHNHTALLYSLGLEQVVDVHQTETNLQ 180
DB 223 KKLGLDFPVDSPVLIQDIASRQLKTRIPKDIHVKVPTLLHSLGMD-LDWKLLQFQ 281
QY 181 SRDGSFLSPASTACVPMHTQNKRCILHFLNVLKFGDVPVCHYPLDLFERLMAVDTV 240
DB 282 CTDGSFLSPSTAYALMQTDHNCNLKNAVHKFNGGVPVVDLFEHIWTVDRLO 341
QY 241 LGIDRYFKKEIKESLDYVRYWDAERGVGWACNPIDVDDTAMGLRIILHGVNVS 300
DB 342 LGISRYFELKIKCIDYFSKYW-TNKGICWARNSPVODIDDTAWAPFLRLHGYAVSAD 400
QY 301 LENFRDEKDFCFAGQTOIGVTDNLNLYRCSQVCPGKEIMEEAKFTTNHLQNALAKN 360
DB 401 FKHF-ESKGEFFCFVQSQNAVTGMVNLRYASHVMSGKILENAKISTSNYLREKRAQN 459
QY 361 NAFDKWAVKDLPGVEYAIKYPWHRSMPLREARSYIEQF-GSDNVWLKTVYKMLYVSN 419
DB 460 QLDDKWIITKDLPGVEYALDVPYASLPRLTRFFLEHYGGDDVWIGKTLRMPVLVN 519
QY 420 EKYLELAKLDFNMVQALHOKETOHIYSVWRRESFNDLTFTRQRPVEMYSVAVSMPEPEF 479
DB 520 SLYLELAKGDYNNCQALHOFWEHRRIRKYYEYECGLRFGSLSEKLLVYTLGSAISFEAQR 579
QY 480 AACRIAYAKTSCLAVIDLDLYDTHGSLD-DLKLFPSEAVRWDISVLDS--VRDNQLKVC 536
DB 580 STERMAWKT--AALMDCVRSFCGSPQVSAAPLCEFAHYSSTALNSRNTEDRLVGVI 636
QY 537 LGLYNTVNGFGKDLKEQGRDVLGYLRKVVWEGLLASYTK---BAEWSAAKYVPTFNEYVE 593
DB 637 LG---TLNHLSTLALLTHGRDIIHRLHAWENWLLTVGEGEGEGGAELIITLN----699
QY 594 NAKVSTALATVNLNLSFFTGELLPDYLIQVDRLSKFLHLSVLTGRLINDTKYQ--AER 651
DB 690 -----LCSVHWISE---EILLS-----HPTYQKLEITNRVSHRLRYKGHSEK 730
QY 652 NRGELVSSVQCYNRENPECTEEBALSHVYGIIDNALKELNWLNPASNAPLCVRILLFN 711
```

```
DB 731 QVGMILTFSE--IEGDMQQLAELVLSH-----SDASELDANIKOTFLT 771
QY 712 TARVMQLFY-MYRDGFGISDKEMKDHVSRITLDPV 745
DB 772 VAK--SFYISAYCD-----DRTINPHIAKVLFRV 799

RESULT 9
Q9MAX2 PRELIMINARY; PRT; 829 AA.
AC Q9MAX2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Copalyl diphosphate.
GN Crotion sublyratus.
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Malpighiales; Euphorbiaceae; Crotonoideae; Crotonaceae;
OC Croton.
OX NCBI_TaxID=107238;
RN [1]
RP SEQUENCE FROM N.A.
RA Sitthithaworn W.; Sankawa U.;
RT "Copalyl diphosphate synthase from Croton sublyratus.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; A8042424; BAA95612.1; -.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth. 1.
DR Pfam; PF03936; Terpene synth. C; 1.
SQ SEQUENCE 829 AA; 96226 MW; 3C3D0C6A848EF0A CRC64;

Query Match 33.6%; Score 1333; DB 10; Length 829;
Best Local Similarity 37.4%; Pred. No. 4.7e-89;
Matches 287; Conservative 155; Mismatches 239; Indels 86; Gaps 16;

QY 1 MFQSGMDGNETPSAYTAWVARIPIIDGSGAPQPTLOWLNNQLPDGSGWBECEIFLAY 60
DB 126 MLKSMDEGEINISAYTAWVRLVKNVNGSGEPFSPALQWISNNQLPDGSGWGDALVENAH 185
QY 61 DRVINTLACLLTLKINWKGDIQVQGVFVRKHEEMKDEADNHRSGFVVPAMLEA 120
DB 186 DRILTSLACVTLKSNWLNHPKVAKGLKFPKENLSKLEENEHEHMSIGFVVPSSLLELA 245
QY 121 KSLGLDLPVHLPIISOIHQKROKLOKIPLVNLHNHTALLYSLGLEQVVDVHQTETNLQ 180
DB 246 RKLDIDVDPDSSHVLOQICECRNTLKKIPKDIHMKMPTSLHSLGEMPD-LEWEKLLKLQ 304
QY 181 SRDGSFLSPASTACVPMHTQNKRCILHFLNVLKFGDVPVCHYPLDLFERLMAVDTV 240
DB 305 FNGSFLSPSTAYALMQTDKCLILYLRIRVQRNGGVNVPVVDLFEHTWAADRLQR 364
QY 241 LGIDRYFKKEIKESLDYVRYWDAERGVGWACNPIDVDDTAMGLRIILHGVNVS 300
DB 365 LGISRLFEQELKECMNVARYW-REDGMCMWARNVDRVDDSSMGFRMLRGLYGHNVSDV 423
QY 301 LENFRDEKDFCFAGQTOIGVTDNLNLYRCSQVCPGKEIMEEAKFTTNHLQNALAKN 360
DB 424 FKFKFKD-DTWVCMPCQSTQAITGMVNLFRASQVLPFGEMILBEAREFCYDFREKOSAN 482
QY 361 NAFDKWAVKDLPGVEYAIKYPWHRSMPLREARSYIEQF-GSDNVWLKTVYKMLYVSN 419
DB 483 AVDDKWIISKDLPGVEYALDVPYASLPVTRFVYIEQYGGDDVWIGKVLRYMPFVN 542
QY 420 EKYLELAKLDFNMVQALHOKETOHIYSVWRRESFNDLTFTRQRPVEMYSVAVSMPEPEF 479
DB 543 NEYLELAKFDYNSCOALHRAEWDNFQWYDECNLGDFGVSKRELLAYIYSAASIFPEK 602
```

```

Qy 480 AACRIAYAKTSCLAVIDDLDTGSLDLDLKFSEAVRRWDISVL-----DSVRDN 530
Db 603 SKERLAWAKTSILLQTDAYPHTNNSIBERKDF---VHQFNMBLLPLHEWNROENGRED 559
Qy 531 QLKVCFLGLYNTVNGFGKGLKEQGRDVLGRLKRVWEGLLASYTKAEWSAAKYVPTFNE 590
Db 660 KAGTCRNTWN-LKRCPLDALVAGHTDISHSRLRHAWEIGLTNWEKGD-----706
Qy 591 YVENAKVSIALATVVLNSIFPTGELLDPYILQOVDLRSKFLHLVSLTGR 640
Db 707 -KRGKAEILVKTINLT---TGPVWSEELL---NCNSKYKQLFSLNLYHLGHFNKN 758
Qy 641 -INDTKTYQAERNRGELVSSVQCYMRENPECTEERALSHVYGIINAKEL-NWELANPA 598
Db 759 KVDDKKKY-----TTPE-----IESSMQELAQVLQNS 787
Qy 699 SNAPLCVRLLFNTARVMQLFMYRDRGFGISDKEMKDHVSRTLDPDV 745
Db 788 DGLDSNIKQTFETV--VKSIIYTTVCDLGTTNY-----HMSKVLFRV 828

RESULT 10
Q9LRB1 ID Q9LRB1 PRELIMINARY; PRT; 827 AA.
AC Q9LRB1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Copalyl diphosphate.
GN CP5.
OS Scoparia dulcis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Plantaginales; Gratiolales; Scoparia.
OX NCBI_TaxID=107240;
RN [1]
RP SEQUENCE FROM N.A.
RA Sitthithaworn W., Sankawa U.;
RT "cDNA encoding Scoparia dulcis CP5."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046689; BAB03594.1;
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth_C; 1.
SQ SEQUENCE 827 AA; 93635 MW; FD94A6C281F9F382 CRC64;

Query Match 33.4%; Score 1327; DB 10; Length 827;
Best Local Similarity 36.7%; Pred. No. 1.3e-88;
Matches 286; Conservative 144; Mismatches 250; Indels 100; Gaps 16;

Qy 1 MFQSMGDETPNSAYDTAWARIPSIDSGAPFPQFTQLWILNNQLPDGSGWGEICFLAY 50
Db 112 MLRSMDEGEITVSAYDTAWALVEDIGSGTGPFPSSLEWISNNQLPDGSGWGDGATFSAH 171
Qy 61 DRVLNTLACLTLKLTWNGDQVQGVFVRKHMEEMKDEADNHRPSGEHVFPAMLDEA 120
Db 172 DRIINTLACVIALRTWNNHSDKSDRGIAFIRENNMYKLEDENEHNPIGFEVALPSLIETA 231
Qy 121 KSLGLDLPVHLFPFISQIHQRKQKLOKPLNVLNHQRTALLYSLRGLQVDVDMQETNLQ 180
Db 232 KKIDIDIPDESGPIFEKRLRKRKFLKRIPRDIIMHKVPTTLHSLSGM-PALDWEKLINQ 290
Qy 181 SRDGSFLSPASTACVFMHTQMKRCLFLNLFVLSKFGDYVPCHPYPLDLFLERLWADTVTR 240
Db 291 SADGSFLSPSTATALQOTKDHNCCLHPARHLQKFGVFNVPVDLFEHLWAVDRLER 350
Qy 241 LGIDRYFKKIKESLDYVRYWDAERGVGWARCNPIPDVDDTAMGLRLRLHLHGVNVSVDV 300

```

```

Db 351 LGLSYFQPEISECIDYVHGHW-TSGICWARSSEVKDIDDITAMGFRLLKLHGYEVSADV 409
Qy 301 LENFRDEKDPFCFAGQTOIGVTDNLNLYRCQVCPGGEKIMBEAKFTTNHLONALAKN 360
Db 410 PKHF-ENGGEPCFVGQSTQAVTGMVNLFRASQVMPFGEDILSNKSKFLSKFLQEKRAAN 468
Qy 361 NAFDQWAKKOLPGEVEYAIKYPWHRSMRPLARSYIEQF-GSNDYWLKGTVTVMYVSN 419
Db 469 ELLDKWIITKOLPGEVYALDVPWYASLPRVETRLYLEQYGGQDDYVIGKTLRYMPYVNN 528
Qy 420 EKVLELAKLDFNMVQALHOKETOHIVSWWRSGFNDLTFTRORPVEMYFSVAVSMPEPEF 479
Db 529 NYILELAKLDYNNCOASHQOEKWSQKYRSCNFGYCHSERSLLLAYIAAASIFEPEP 588
Qy 480 AACRIAYAKTSCLAVIDDLDTGSLDLDLKFSEAVRRWDISVLDSVRDNQKVCFLGL 539
Db 589 ALERLAWAKTALIMETIETILKDNSSQE-----SGNVLSMNPMLMAAACM-- 634
Qy 540 YNTVNGFGKGLKEQGRDVLGRLKRVWEGLLASYTKAE-----WSNA----- 582
Db 635 -----QTEQVAGITKQ--ERLTETLVRASKPELLIGTRCMTAGTSINCM 678
Qy 583 ---KVPTFNEY-VENAKVSIALATVVLNSIFPTGELLDPYILQOVDLRSKFLHLVSLTG 638
Db 679 GRKSGRLTCEERDVHSGQDAELLVRLNLCGGGRWSEELLSS---HPKYHLLLEVTVN 735
Qy 639 RLINDTKTYQAERN-----RGELVSSVQCYMRENPECTEERALSHVYGIIDNAL 687
Db 736 RVCHQLRLFRQACDVNGCGMTDVGILSAQIESNMQELAKLVPTKSSS---GLDSDT 792
Qy 688 KEINWELANPASNAPLCVRLLFNTARVMQLFY--MYRDRGFGISDKEMKDHVSRTLDPDV 745
Db 793 KQ-----NFLTARSFYAYV-----CNPGSINFHIAKVLFRV 826

RESULT 11
Q38802 ID Q38802 PRELIMINARY; PRT; 802 AA.
AC Q38802;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE ENT-KAURENE synthetase A.
GN GAL OR T5J8.9 OR AT4G02780.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
EX MEDLINE=95086384; PubMed=7994182;
RA Sun T.P., Kamiya Y.;
RT "The Arabidopsis GAL locus encodes the cyclase ent-kaurene synthetase
A of gibberellin biosynthesis."
RL Plant Cell 6:1509-1518(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cnv. Columbia;
RA de la Bastide M., Gnoj L., Habermann K., Huang E.N., Gottesman T.,
RA Kaplan N., Lodhi M., Jensen K., Hameed A., Schutz K., Martienssen R.,
RA Dedhia N., Parnell L.D., McCombie W.R.;
RT "Arabidopsis thaliana BAC T5J8 from chromosome IV, short arm."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Spiegel L.A., Huang E.N., Nascimento L.U., de la Bastide M., Vil D.M.,
RA Preston R.R., Matero A., Shah R., O'Shaughnessy A., Rodriguez M.,
RA Shekher M., Schutz K., See L.H., Swaby I., Habermann K., Dedhia N.N.,
RA Newes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

```





```

QY 532 LKVCFLGYNTVNGFGKDLKEQGRDVLGYLRKRWEGILLASYTK---EAEWSAAKYVPTF 588
Db 632 WHGLMVALKGLTHETALDVLMTTHRRDIHPQLHAWEMLMRWQGVDAITEGQAEILVQRTI 591
QY 589 N----EYVEN---AKVSTALATVLSNI 609
Db 692 NWTAGRWVSNEALLAHPOYLLSSVNNI 719

RESULT 13
Q9ZTN8 PRELIMINARY; PRT; 823 AA.
AC Q9ZTN8;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Copalyl diphosphate synthase 1.
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063778; PubMed=9847116;
RA Smith M.W., Yamaguchi S., Ait-Ali T., Kamiya Y.;
RT "The first step of gibberellin biosynthesis in pumpkin is catalyzed by
RT at least two copalyl diphosphate synthases encoded by differentially
RT regulated genes.";
RL Plant Physiol. 118:1411-1419 (1998).
DR EMBL; AF049905; AAD04292.1; -.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
DR SEQUENCE 823 AA; 93946 MW; 3F367AAF474721F7 CRC64;

```

```

Query Match 32.6%; Score 1296.5; DB 10; Length 823;
Best Local Similarity 38.3%; Pred. No. 2.2e-86;
Matches 293; Conservative 135; Mismatches 271; Indels 67; Gaps 19;

QY 1 MFQSMGDETNPSAYDTAWARIPSI-DGSGAPQPPQTLOWLNNQLPDGSGEGECIFLA 59
Db 103 MLSSMDGDISISAYDTAWALIPRVLDGKTPLEPSSLEWIAQNLQPDGSGEGSIFSA 162

QY 60 YDRVLTALCLLTLLKINWKGDIQVQGVFVRKHEEMKDEADNHRPSGFVFPFAMLDE 119
Db 163 HDRILSTLACVLCALNSKWLHPDKSEKGMVFLNKNISKLEDAEHAHMLIGFEVAPPSLMF 222

QY 120 AKSLGLDLPHYLPFTSIQHQKQKLOKIPNLVNHQTLALLYSLEGQDV--VDMQBIT 177
Db 223 AKRLNLQVPTDSPVLQEIHNRRSIKLIPKEIMHKVPTTLHSLGMEGLDGMGLL 282

QY 178 NLQSRDGSFLSPASTACVPHHTQMKRCLHFLNLFVLSKFGDVPCHPYLDLPERLWAVDT 237
Db 283 KLQAPDGSFLKSPASTAFAPKTNNSCFKYLVSVRFGVGNVPYVVDLFEHIAVDR 342

QY 238 VERLIGIDRYPKKIKESLDYVRYWDAERGVGWARCNPIPDVDDTAMGLRIILRLHGYNS 297
Db 343 LQRLGVSRFFPHPIEVESVDLRRHW-TDKGICWADVDFYDIDDTAMGPKLLRLFGHEVS 401

QY 298 SDVLENFRDEKDFPCFAGQTQIGTNDNLNLYRCS-QVCFPEKIMEAKTITTHLQNA 356
Db 402 AEVFKNP-EKDGFEVCIAGQSQAVTGMFNLYRASDVNMFPGKEKILEDAKQFSYKFLREK 460

QY 357 LAKNAPDKKAVKQDLPGEVEAIVKYPWHRSPLEARSYIQF-GSNDVWLKGTVYKML 415
Db 461 QAADELLDKWIIITKDLPEGVGVALDVPWFASLPRVETRYFIBQYCGENDIWKRTLYRMF 520

```

```

QY 416 YYSNEKYLEAKLDFNMVQALLHOKETQHIVSWRESGFNDLITTRQRPVEMYSVASVMP 475
Db 521 KYNNPTYLEAKLDYNNKQLLHQNEWVDIQKWTYNNLRDYGMRRTSLLSYFGAACSIIF 580
QY 476 EPEFAACRIAYAKTSCILAVILDDLDYTHGSLDDKLKLPSEAVRRWDISVLDSVRDN---- 530
Db 581 EPERAKERLAWTKTAALVCAIESHP-KDANADORRAFIQOPIINF--AIDQADYTNWRA 637
QY 531 ---QLKVCFLG----LYNTVNGFGKDLKEQGRDVLGYLRKRWEGILLASYTKAEWSAAK 583
Db 638 GNVQKGGGQGLVGIILLRTLSISLILVSHGFDITHHLHQAWEKWLFKWOEDGD----- 692
QY 584 YVPTFNEYVENAKVSTALATVLSNIFFTCGELLPDYLQOVDLRS--KPLHLVSLTGLLI 641
Db 693 -----VHKEAEALLVGTIILNSGCSLLE-----DLLSNPQYQKLSYLTNKCVC 734
QY 642 NDTKTYQAEK--NRGELSVSVOCYMRNEPECTEERALSHPVYGIIDNALKEINWELANPAS 699
Db 735 HQGHFKGHKVTNGGLYKKTENKM---PPEIEDMRKLLQWVIONS-SQGN-DIDSPK 789
QY 700 NAPLCVRRLLFNTARVMQLFYMYRDGFGISDKEMKDHVSRTLPDPV 745
Db 790 NTFLLTVAKSSVYAA-----YFDPMTIN-----YHIAKVLPERV 822

```

## RESULT 14

```

Q84LM0 PRELIMINARY; PRT; 822 AA.
AC Q84LM0;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Copalyl diphosphate synthase.
GN CMCP51.
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1]
RP SEQUENCE FROM N.A.
RA Igasaki T.;
RT "Cucurbita maxima mRNA for copalyl diphosphate synthase 1, complete
RT cds.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB109763; BAC76429.1; -.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR005630; Terpene synth C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp_cyc Toroid.
DR InterPro; IPR001906; Terp_synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth C; 1.
SQ SEQUENCE 822 AA; 93778 MW; 7BFA086EFA7BE0D5 CRC64;

```

```

Query Match 32.6%; Score 1294.5; DB 10; Length 822;
Best Local Similarity 38.3%; Pred. No. 3.1e-86;
Matches 293; Conservative 134; Mismatches 272; Indels 67; Gaps 19;

QY 1 MFQSMGDETNPSAYDTAWARIPSI-DGSGAPQPPQTLOWLNNQLPDGSGEGECIFLA 59
Db 102 MLSSMDGDISISAYDTAWALIPRVLDGKTPLEPSSLEWIAQNLQPDGSGEGSIFSA 161

QY 60 YDRVLTALCLLTLLKINWKGDIQVQGVFVRKHEEMKDEADNHRPSGFVFPFAMLDE 119
Db 162 HDRILSTLACVLCALNSKWLHPDKSEKGMVFLNKNISKLEDAEHAHMLIGFEVAPPSLMF 221

QY 120 AKSLGLDLPHYLPFTSIQHQKQKLOKIPNLVNHQTLALLYSLEGQDV--VDMQBIT 177
Db 222 AKRLNLQVPTDSPVLQEIHNRRSIKLIPKEIMHKVPTTLHSLGMEGLDGMGLL 281

```

```
QY 178 NLSRDSGSLSSPASTACVFMHTONKRCFLHFLNVLKFGDYVYPCHYPDLDFERLWAVDT 237
Db 282 KLAQDGSFLKSPASTAFAPKATNSNCFKYLSVSRFNGVNPVVDLFEHIAVDR 341
QY 238 VERLGIDRYPKGKIKESLDYVRYWDAERGVMGWCNPIPDVDDTAMGLRILRLHGYNVS 297
Db 342 LQRLGVSRRPPEBIVEVDYLRHHW-TDKGICWARDVEFYDIDDITAMGFKLLRFGHEVS 400
QY 298 SDVLENFRDEKGFPCFAGTQIGVTDNLNLYRCS-QVCPGGEKIMEEAKFTTNHLQNA 356
Db 401 AEVFNK-EXDGEFVCIAGOSTQAVTFNLYRASDQVMPFGEKILEDAKQFSYKFLREK 459
QY 357 LAKNAPDKWAKVADLPGEVEYAIKYPWHRSMRLEARSIEOF-GSDVMWLGKTVYKML 415
Db 460 QADELLDKWIIYKDLPEGEVGYALDVWPFASLPRVETRYFIEQYGGENDIWIGKTYRMF 519
QY 416 YVNEKYLELAKLDFNMVQALHOKETOHIYVSWRESGFNDLTFTQRPVEMVYSVAVSMF 475
Db 520 KVNNDTYLELAKLDYKQCLLHQNWEVDIQKWTENNRLDYGMRRTSLPSYFGAACSIF 579
QY 476 EPEFAACRIAYAKTSCLAVIDLDLYDTHGSLDLDLKLSEAVRWDISVLSVRDN----- 530
Db 580 EPERAKERLAWTKTAALVGAIESHF-KDANADQRRAFIQFINFD--AIDQAYDTNAWRA 636
QY 531 ---QLKVCFLG---LYNTVNGFGKGLKEQGRDVLGRLKRVWEGLIASYTEKAEWSAAK 583
Db 637 GNVOQGGGGLVILLRTLSLSLDLVSHGFDITHHLQAEKWLFKWQEDG----- 691
QY 584 YVPTFNEVENAKVSIATVNLNSIFFTGELLPDYILOQVDLRS--KFLHLVSLTGRLI 641
Db 692 -----VHKEBAELVGTIILNSGSGSTLE-----DLLSNPOYQKLSYLTNKVC 733
QY 642 NDKTYQAE--NRGELVSSVQCYMRENPECTBEALSHVYGIIDNALKEINELANPAS 699
Db 734 HQLGFKKHVKTNGGIYKTEKNKM-----PPEIBEDMKLMQMIQNS--SDGN-DIDSPK 788
QY 700 NAPLCVRLFLNTARVMQVFMVYRDGFGISDKEMKDHVSRTLDPV 745
Db 789 NTELTVAKSYYAA-----YFDPWTIN-----YHIAKVLPERV 821

RESULT 15
Q41771 PRELIMINARY; PRT; 823 AA.
AC Q41771;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Kaurene synthase A.
GN AN1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95210929; PubMed=7696880;
RA Bensen R.J., Jhal G.S., Crane V.C., Tossberg J.T., Schnable P.S.,
RA Mealey R.B., Briggs S.P.;
RT "Cloning and characterization of the maize An1 gene.";
RL Plant Cell 7:75-84 (1995).
DR EMBL: L37750; AAA73960.1; -.
DR F01397; T02959.
DR GO: 0016829; F1-yase activity; IEA.
DR GO: 0008152; P:metabolism; IEA.
DR InterPro: IPR005630; Terpene synth. C.
DR InterPro: IPR008949; Terpenoid synth.
DR InterPro: IPR008930; Terp. cyc. toroid.
DR InterPro: IPR001906; Terp_synth-like.
DR Pfam: PF01397; Terpene synth; 1.
DR Pfam: PF03936; Terpene synth. C; 1.
SQ SEQUENCE 823 AA; 95123 MW; E4D9C97CDB7A0B34 CRC64;
```

```
Query Match 32.6%; Score 1293.5; DB 10; Length 823;
Best Local Similarity 36.5%; Pred. No. 3.7e-86;
Matches 289; Conservative 146; Mismatches 243; Indels 113; Gaps 24;

QY 1 MFQSGMDGETNPSTAYTAWARIPSI--DGSAGAPQFPQTQLQWLNQPLDGSMBEECIFL 58
Db 100 MLRSMNDGDISASAYDTAWAMVPKVGDDGAQFPATVRWIVDHLQPLDGSMDGSALFS 159
QY 59 AYDRVLTNLAULTLKIWKNGDIOVKQGVFVRKHEEMKDEADNHRPSPSEVVFVPMALD 118
Db 160 AYDRMINTLACVVALLTKWLEPARCEAGLSFLENMWRLEAEBAESMPICEAFAPSLIQ 219
QY 119 EAKSLG-LDLPYHLPISTQIHOKRQKLOKIPLNVLNHNHOTALLYSLEGQDVVDWQBEIT 177
Db 220 TARDLGVDFFPYGHALQSIYANREVAKRIPDMHHRVFTSLHSELEGNPD-LDWFRLL 278
QY 178 NLSRDSGSLSSPASTACVPMHTONKRCFLHFLNVLKFGDYVYPCHYPDLDFERLWAVDT 237
Db 279 NLQSCDGSFLSPSATAYALMQTDKKCPYIDRIIVKKGNGVNPVVDLFEHIAVDR 338
QY 238 VERLGIDRYPKKIKESLDYVRYWDAERGVMGWCNPIPDVDDTAMGLRILRLHGYNVS 297
Db 339 LERLGISRYFQREIEQCMDYVNRHW-TEDGICWARKSNVKDVEDDTAWAFRLRLHGYNVS 397
QY 298 SDVLENFRDEKGFPCFAGTQIGVTDNLNLYRCSQVCPGGEKIMEEAKFTTNHLQNAL 357
Db 398 PSVFNK-EXDGEFVCIAGOSTQAVTFNLYRASDQVMPFGEKILEDAKQFSYKFLRE 456
QY 358 AKNAPDKWAKVADLPGEVEYAIKYPWHRSMRLEARSIEOF-GSDVMWLGKTVYKML 416
Db 457 EQGMIRDKNIVAKDLPEGEVQYTLDFPWAYSLPRVEARTYLDQYGGKDDVMIGKTYRMPL 516
QY 417 VNEKYLELAKLDFNMVQALHOKETOHIYVSWRESGFNDLTFTQRPVEMVYSVAVSMFE 476
Db 517 VNNDTYLELAIIRDNFHCQALHQCNGLOTWYKDNCLDAFGVBPQDVLRSYFLAAACIFE 576
QY 477 EPEFAACRIAYAKTSCLAVIDLDLYDTHGSLDLDLKLSEAVRWDISVLSVRDNQNLK-- 533
Db 577 PSRAAERLAWARTSMIA-----NAISTH-----LRDISDKKRLCECFVHCLYEENDVSWLK 627
QY 534 -----VCFGLYNTVNGFGKGL--KEQGRDVLGRLKRVWEGLIASYTEKAEWSAAKYV 585
Db 628 RNPNDVILERALRLINLLAQEALPTHEGQRFTHSLLSLAW-----TEMMLQKAN 677
QY 586 PTNEVEYENAKVSIATVNLNSIFFTGELLPDYILOQVDLRSKFLHLVSL-----TGRLI 641
Db 678 KEENKYHKCSGIE-----POTMVHD-----RQTYLLLVQVIEICAGRI- 715
QY 642 NDKTYQAEERNRGELVSSVQCYMRENPE-----ECTEEALSHVYGIIDNALK-----ELNW- 692
Db 716 -----GEAVSMIN--NKDNDWFITLQCATCDLSNHRMLLSQDTMKNEARINWI 761
QY 693 -----ELA-----NPASNAPLCVRLFLNTARVMQVFMVYRDGFGISDKEMK 735
Db 762 EKEIELNMQELAQSLLRCDKETSNNK--TKTKLMDVLR--SLYYATH-----SPOHMID 812
QY 736 -HVSRTLDPV 745
Db 813 RHVSRRVIFEPV 823

Search completed: August 17, 2004, 21:02:30
Job time : 49.0877 secs
```